National Library
of Canada

Acquisitions and
Bibliographic Services Branch

395 Wellington Street
Ottawa, Ontario
K1A 0N4

Bibliothèque nationale
du Canada

Direction des acquisitions et
des services bibliographiques

395, rue Wellington
Ottawa (Ontario)
K1A 0N4

Your file Votre référence

Our file Notre référence

NOTICE

The quality of this microform is heavily dependent upon the quality of the original thesis submitted for microfilming. Every effort has been made to ensure the highest quality of reproduction possible.

If pages are missing, contact the university which granted the degree.

Some pages may have indistinct print especially if the original pages were typed with a poor typewriter ribbon or if the university sent us an inferior photocopy.

Reproduction in full or in part of this microform is governed by the Canadian Copyright Act, R.S.C. 1970, c. C-30, and subsequent amendments.

AVIS

La qualité de cette microforme dépend grandement de la qualité de la thèse soumise au microfilmage. Nous avons tout fait pour assurer une qualité supérieure de reproduction.

S'il manque des pages, veuillez communiquer avec l'université qui a conféré le grade.

La qualité d'impression de certaines pages peut laisser à désirer, surtout si les pages originales ont été dactylographiées à l'aide d'un ruban usé ou si l'université nous a fait parvenir une photocopie de qualité inférieure.

La reproduction, même partielle, de cette microforme est soumise à la Loi canadienne sur le droit d'auteur, SRC 1970, c. C-30, et ses amendements subséquents.

Canada

UNIVERSITY OF ALBERTA

VALIDATION OF THE CANADIAN OCCUPATIONAL PERFORMANCE MEASURE

BY

CHETWYN CHE HIN CHAN ©

A THESIS SUBMITTED TO THE FACULTY OF GRADUATE
STUDIES AND RESEARCH IN PARTIAL FULFILMENT OF THE
REQUIREMENTS OF THE DEGREE OF
DOCTOR OF PHILOSOPHY

EDUCATIONAL PSYCHOLOGY

EDMONTON, ALBERTA

FALL, 1995

National Library
of Canada

Acquisitions and
Bibliographic Services Branch

395 Wellington Street
Ottawa, Ontario
K1A 0N4

Bibliothèque nationale
du Canada

Direction des acquisitions et
des services bibliographiques

395, rue Wellington
Ottawa (Ontario)
K1A 0N4

Your file Votre référence

Our file Notre référence

THE AUTHOR HAS GRANTED AN IRREVOCABLE NON-EXCLUSIVE LICENCE ALLOWING THE NATIONAL LIBRARY OF CANADA TO REPRODUCE, LOAN, DISTRIBUTE OR SELL COPIES OF HIS/HER THESIS BY ANY MEANS AND IN ANY FORM OR FORMAT, MAKING THIS THESIS AVAILABLE TO INTERESTED PERSONS.

L'AUTEUR A ACCORDE UNE LICENCE IRREVOCABLE ET NON EXCLUSIVE PERMETTANT A LA BIBLIOTHEQUE NATIONALE DU CANADA DE REPRODUIRE, PRETER, DISTRIBUER OU VENDRE DES COPIES DE SA THESE DE QUELQUE MANIERE ET SOUS QUELQUE FORME QUE CE SOIT POUR METTRE DES EXEMPLAIRES DE CETTE THESE A LA DISPOSITION DES PERSONNE INTERESSEES.

THE AUTHOR RETAINS OWNERSHIP OF THE COPYRIGHT IN HIS/HER THESIS. NEITHER THE THESIS NOR SUBSTANTIAL EXTRACTS FROM IT MAY BE PRINTED OR OTHERWISE REPRODUCED WITHOUT HIS/HER PERMISSION.

L'AUTEUR CONSERVE LA PROPRIETE DU DROIT D'AUTEUR QUI PROTEGE SA THESE. NI LA THESE NI DES EXTRAITS SUBSTANTIELS DE CELLE-CI NE DOIVENT ETRE IMPRIMES OU AUTREMENT REPRODUITS SANS SON AUTORISATION.

ISBN 0-612-06191-4

Canada

Department of Rehabilitation Sciences
The Hong Kong Polytechnic University
Hung Hom Kowloon
Hong Kong

The Publisher
American Council on Education
One Dupont Circle NW
Suite 800
Washington
DC 20036-1193
USA

4th September 1995

Dear Publisher,

Permission to Copy

I am a doctoral student at the Department of Educational Psychology, University of Alberta, Canada. I am presently undertaking my research on the validation of the Canadian Occupational Performance Measure (COPM). I would like to include the following material in my doctoral dissertation titled "Validation of the Canadian Occupational Performance Measure".

Name of Author:	Messick, S. (1989) in R.L. Linn (Ed.).
Name of Publication:	Educational Measurement (3rd.ed.)., p.50
Title of Figure:	Figure 2.4 Components of Construct Validity

I would appreciate if you would sign below to indicate your permission for me to copy the stated material for this purpose. Thank you.

Sincerely,

Chetwyn Chan, MSc.,OT(C).

I give my permission to Chetwyn Chan to copy Figure 2.4 on page 50 of Educational Measurement (3rd.ed.). (1989) and incorporate it in his doctoral dissertation titled "Validation of the Canadian Occupational Performance Measure".

Publisher, American Council on Education

9/7/95
Date

UNIVERSITY OF ALBERTA

RELEASE FORM

NAME OF AUTHOR: Chetwyn Che Hin Chan

TITLE OF THESIS: Validation of the Canadian Occupational Performance Measure

DEGREE: Doctor of Philosophy

YEAR THIS DEGREE GRANTED: 1995

Permission is hereby granted to the University of Alberta Library to reproduce single copies of this thesis and to lend or sell such copies for private, scholarly or scientific research purposes only.

The author reserves all other publication and other rights in association with the copyright in the thesis, and except as hereinbefore provided, neither the thesis nor any substantial portion thereof may be printed or otherwise reproduced in any material form whatever without the author's prior written permission.

Flat 1204, Tai Hang Terrace,
5, Chun Fai Road, Hong Kong

September 21, 1995

UNIVERSITY OF ALBERTA

FACULTY OF GRADUATE STUDIES AND RESEARCH

The undersigned certify that they have read, and recommend to the Faculty of Graduate Studies and Research for acceptance, a thesis entitled VALIDATION OF THE CANADIAN OCCUPATIONAL PERFORMANCE MEASURE submitted by CHETWYN CHE HIN CHAN in partial fulfilment of the requirements for the degree of DOCTOR OF PHILOSOPHY.

TMaguire
Dr. T. O. Maguire

TRogers
Dr. T. Rogers

S. Hunka
Dr. S. Hunka

Dr. H. Madill

Dr. D. K. Kieren

P. Nagy
Dr. P. Nagy

Date: 18th September, 1995

DEDICATION

I dedicate this work to my wife Tatia Lee who has supported me throughout the entire course of my doctoral study. I also dedicate this work to my friends who I met and enjoyed during the time in Edmonton.

ABSTRACT

The Canadian Occupational Performance Measure (COPM) is used to assess clients' perceived performance and satisfaction with their performance of daily activities. The purpose of this study was to gather evidence on the substantive, content, structural, and criterion-related validity with reference to Messick's validation model. In addition, the test-retest stability and other procedural and utility properties of the COPM were studied.

Thirty-nine adult clients from two diagnostic groups (orthopedic and stroke) at the Glenrose Rehabilitation Hospital in Edmonton, Canada participated in the study. Clients were assessed at the initial and pre-discharge assessment occasions using the COPM as well as the Role Checklist, Klein-Bell ADL Scale, Satisfaction with Performance Scaled Questionnaire (SPSQ), and the Functional Independence Measure (FIM). Nine clinical and measurement experts participated in an expert panel review to evaluate the test content, the scoring system and other psychometric properties of the COPM.

Results from the expert panel review suggested that the "Performance" and "Satisfaction" subscale scores reflected clients' occupational performance, but not the performance components. Substantive-related validity was supported with the findings of consistency between the activities identified by the clients in the COPM and their perceived life roles. The testing process was found to be relevant to a client-centred model of practice. Analysis of clients' protocols revealed problems in using the self-report 10-point rating scale to measure the "importance" and "satisfaction" constructs. Protocol analysis also revealed the differences among the clients in interpreting and understanding those constructs. Criterion-related validity of the COPM was demonstrated by the correlations between the COPM subscale scores and the Motor subscale of the FIM (r=-.14 to .38) and the SPSQ (r=-.13 to .40). The stability estimates of the COPM varied from .32 for the Performance subscale and .09 for the Satisfaction subscale. Evidence gathered in this study indicated the need for further revisions of the COPM to improve the accuracy and meaning of its subscale and "change" scores. This study also demonstrated the importance of vigorously testing the meaningfulness and truthfulness of clinical outcome measures used in health care services.

ACKNOWLEDGMENT

I am indebted to a number of people who were supportive and helpful throughout the program of my study. I would like to thank the members of my thesis committee, especially Dr. T.O. Maguire, my project supervisor, who has guided me to make wise decisions both in my study, research and career. I wish to extend my gratitude to Dr. T. Rogers and Dr. S. Hunka whom I have admired since I became a CRAMER for their intriguing insights into psychometric, measurement and statistics. Special thanks also go to Dr. H. Madill and Dr. D. Kieren for their support and useful suggestions. In addition to my committee members, I wish to extend a special thanks to all the staff at the Occupational Therapy Department, Glenrose Rehabilitation Hospital, Dr. L. Liu, Ms. Annette Rivard and Dr. J. Edwards who participated in the data collection process. Last but not the least, I would like to thank Dr. K. Pain and the staff of the Research Services Department, Glenrose Rehabilitation Hospital for the administration work and granting of research funding to support this research.

This research was funded by the Research Fund, Glenrose Rehabilitation Hospital. However, it does not necessarily support the findings and conclusions of the research.

TABLE OF CONTENTS

Chapters Page

I. INTRODUCTION 1

Statement of the Purpose 1

Background and Justification of the Study 1

Limitations of the Study 6

Organization of the Chapters 7

II. LITERATURE REVIEW 9

Introduction 9

The Model of Occupational Performance 9

Canadian Occupational Performance Measure (COPM) 25

Validation Process in Occupational Therapy 35

Messick's Concept of Validity - Toward a Validation Model 38

Strategies of Establishing Evidence of Construct Validity 49

Conclusion 56

III. METHOD OF INVESTIGATION 58

Introduction 58

Sample and Sampling Method 58

Instrumentation 61

Design of the Validation Process 73

Chapters Page

Stability of the COPM 75

Data Collection Procedures 78

IV. RESULTS AND BACKGROUND ANALYSIS 87

Introduction 87

Clients and Assessment Results 87

Results of COPM 89

Results of Satisfaction with Performance Scaled Questionnaire 98

Results of Klein-Bell ADL Scale 101

Results of Functional Independence Measure 103

Conclusion 106

V. EVIDENCE OF CONSTRUCT VALIDITY OF COPM 107

Introduction 107

Section A - Content-Related Evidence 107

Section B - Substantive-Related Evidence 127

Section C - Structural Related Evidence 173

Section D - Criterion-Related Evidence 215

VI OTHER PSYCHOMETRIC PROPERTIES 231

Introduction 231

The Utility and Procedural Review 231

Chapters Page

Evaluation of Testing Procedures of COPM 232

Evaluation of Standardization of COPM 234

Evaluation of Scientific Rigor of COPM 236

Evaluation of Clinical Utility of COPM 238

Evaluation of Control of Testers' Competence 240

Conclusion on Utility and Procedural Evaluation 242

VII. CONCLUSION AND SUGGESTIONS FOR CHANGES 244

Introduction 244

COPM Encompasses Three Areas of Occupational Performance 245

COPM as a Measure of Performance Component 248

COPM Incorporates Roles and Role Expectations of Clients .. 251

COPM Incorporates Importance of Activities of Clients 253

COPM Measures Client's Self-Perceived Performance 256

COPM Measures Client's Satisfaction with Performance 258

Use of Ten-Point Rating Scale 259

Scoring System of COPM 261

Stability of COPM 271

Universal Utility of COPM 272

Chapters Page

Conclusion 272

REFERENCES 275

APPENDICES 288

LIST OF TABLES

Table Page

3.1 Assessment Protocol of the Clients in the Three Assessment Occasions 83

4.1 Mean Age of Clients in the Orthopedic and Stroke Groups 89

4.2 Types of Disability of Clients used in the COPM 91

4.3 Subscale Scores on the COPM of Clients 93

4.4 Mean Performance and Satisfaction Change Scores of Clients 98

4.5 Mean SPSQ Home Management Scores (HMS) of Clients 99

4.6 Mean SPSQ Social/Community Problem Solving Scores (SCS) of Clients 100

4.7 Mean Klein-Bell ADL Scale Scores of Clients 102

4.8 Mean FIM Motor Scores of Clients 104

4.9 Mean FIM Cognitive Scores of Clients 105

5.1 Relevance of COPM to Assessment of Occupational Performance 110

5.2 Relevance of COPM to Assessment of Performance Components 114

5.3 Overall Relevance and Representativeness of COPM to Model of Occupational Performance 123

Table Page

5.4 Frequencies of Activity Categories Identified by the Clients in the COPM 130

5.5 Eleven Activities Most Frequently Identified as a "Concern" in the COPM 132

5.6 Clients' Importance Ratings of the Activities Most Frequently Identified in COPM 135

5.7 Comparisons of Frequencies of Type of Activities Selected by Orthopedic and Stroke Groups 139

5.8 Comparison of the Proportions of Clients Perceived Future Life Roles between Different Gender and Diagnostic Groups 142

5.9 Patterns of Activity and Role of Student 144

5.10 Patterns of Activity and Role of Worker 146

5.11 Patterns of Activity and Role of Care Giver 147

5.12 Patterns of Activity and Role of Home Maintainer 148

5.13 Patterns of Activity and Role of Hobbyist/Amateur 149

5.14 Comparison of the COPM Mean Performance and Satisfaction Scores between Initial Assessment and Reassessment 176

5.15 Pattern of Activities Selected by Clients between Initial Assessment and Reassessment 180

Table Page

5.16 Activities Identified by One Client in Initial Assessment and Reassessment 183

5.17 Activities Identified by Another Client in Initial Assessment and Reassessment 185

5.18 Inter-item Correlation of Importance Ratings of COPM's Activity Items 189

5.19 Inter-item Correlation of Performance Scores of COPM's Activity Items in both Initial and Pre-discharge Assessments 192

5.20 Inter-item Correlation of Satisfaction Scores of COPM's Activity Items in both Initial and Pre-discharge Assessments 196

5.21 Correlation of Initial Item Performance and Satisfaction Scores on the COPM 199

5.22 Initial Mean Performance and Satisfaction Ratings of Activity Items 200

5.23 Correlations of Pre-discharge Item Performance and Satisfaction Scores on the COPM 201

5.24 Pre-discharge Mean Performance and Satisfaction Ratings of Each Activity Item 202

5.25 Correlation of Change in Performance and Satisfaction Ratings on the COPM 203

Table Page

5.26 Correlation of COPM Performance and Satisfaction Subscale Scores in Initial and Pre-discharge Assessments 205

5.27 Structural Fidelity of Using the 10-point Rating Scale in the COPM 209

5.28 Structural Fidelity of Computation of Item and Total Scores on the COPM 211

5.29 Expected Convergent and Discriminant Validity between COPM and KB-ADL Scale 216

5.30 Convergent and Discriminant Validity between COPM and KB-ADL Scale for Initial and Pre-discharge Assessment 218

5.31 Expected Convergent and Discriminant Validity between COPM and SPSQ 220

5.32 Convergent and Discriminant Validity between the COPM and SPSQ for Initial and Pre-discharge Assessments 222

5.33 Expected Convergent and Discriminant Validity between COPM and FIM 225

5.34 Convergent and Discriminant Validity between COPM and FIM for Initial and Pre-discharge Assessment 226

6.1 Expert Panel Review on Testing Procedures of COPM 233

Table Page

6.2 Expert Panel Review Evaluation of Standardization of COPM 235

6.3 Expert Panel Review on Scientific Rigor of COPM 237

6.4 Expert Panel Review Evaluation of Clinical Utility of COPM 239

6.5 Expert Panel Review Evaluation of Testers' Competence using COPM 242

7.1 Relative Importance Index of the COPM 264

7.2 Item Performance Scores Using Importance Rating and Relative Importance Index 265

LIST OF FIGURES

Figure

Page

2.1 Messick's Facets of Unified Validity 40

2.2 Messick's Components of Construct Validity 43

5.1 Scatterplot of the COPM Initial and Reassessment Performance
Subscale Scores 177

5.2 Scatterplot of the COPM Initial and Reassessment Performance
Subscale Scores 178

LIST OF APPENDICES

Appendix Page

I Glossary of Terms - Model of Occupational Performance 288

II Sample Items of the Role Checklist 290

III Sample Items of the SPSQ 292

IV Sample Items of the Klein-Bell ADL Scale 293

V Sample Items of the Functional Independence Measure 294

VI Panel Review Questionnaire 295

VII Educational Session 311

VIII Consent Form 315

IX Client Information Sheet 317

X Client Data Base Sheet 318

XI Activity Reference List 319

CHAPTER I

INTRODUCTION

The purpose of Chapter I is to orientate readers to the context of the use and validation of clinical assessments in the field of occupational therapy. It begins with a purpose statement which summarizes the objectives of the present research study. This is followed by a section on the background and justification of the study, limitations of the study, and, finally, an introduction to the content and organization of the dissertation.

Statement of Purpose

The Canadian Occupational Performance Measure (COPM) (Law, Baptiste, Carswell-Opzoomer, McColl, Polatajko, & Pollock, 1991) is an assessment instrument built on the model of occupational performance used by occupational therapists. The purpose of this research was to gather evidence of construct validity of the use and interpretation of the COPM in its measurement of clients' occupational performance who underwent rehabilitation services in a hospital in Canada.

Background and Justification of the Study

In the past decade, more and more pressure has been put on rationalizing the costs of health care services. The reality of a poor economy and the huge deficits that governments are facing have forced health service providers and politicians to conduct ongoing reviews of their health care expenditures. Other factors such as an increasing prevalence of chronic disease in an aging population, proliferation of technological advances in health care, and an increase in health care recipients as discerning consumers, have in turn led to increased demands for high quality services. As a consequence, the demand for sound clinical outcome measures has increased with the demands of determining "effectiveness" and "worthiness" of clinical programs in the face of declining resources.

In the United States, much effort has been put into the creation of a centralized data system to measure medical and rehabilitation outcomes. In 1983, the National Task Force was established to develop the Uniform Data System for Medical Rehabilitation (Center for Functional Assessment Research, 1991). The task force produced an instrument called the Functional Independence Measure (FIM). The FIM was constructed and standardized as the principal measure of outcome of care in terms of the level of independence or degree of severity relevant to medical rehabilitation (Hamilton, Granger, Sherwin, Zielezny, & Tashman, 1987). It measures "a selected minimum number of key activities intended to be necessary and sufficient indicators of level or cost of disability" (p.141). In order to quantify the outcomes of medical rehabilitation services, data are collected from numerous facilities across the United States and are managed and analyzed with sophisticated computer software. Recently, the Task Force employed the Rasch model to assist data analysis and numerous articles have been published on the results of using this procedure (Granger, Hamilton, Linacre, Heinemann, & Wright, 1993; Heinemann, Linacre, Wright, Hamilton, & Granger, 1993).

In Canada, however, no global nation-wide clinical outcome measures have been developed that can serve as a counterpart to the FIM. Instead, these initiatives have been taken up by local governing bodies, professional organizations, or individual facilities. Therefore, there is no collective development of a universal clinical outcome measure in this country.

In occupational therapy, the use and development of standardized assessment tools and valid outcome measures are emphasized throughout therapists' education and professional practice (Campbell, 1989; Christiansen, 1991b, 1993; Law, 1987, 1993). Many pioneers in the field have developed and defined the construct of "occupational performance" upon which clinical assessment tools and outcome measures have been and are continuing to be developed (CAOT, 1991; Christiansen, 1991b; Christiansen & Baum, 1991; Law, 1993; Law et al., 1991; Mathiowetz, 1993; Pollock, Baptiste, Law, McColl, Opzoomer, & Polatajko, 1990; Reed & Sanderson, 1980; Trombly, 1993). In dealing with clients with physical disabilities, instead of solely measuring physical performance components such as strength, range of motion, muscle tone and motor coordination, occupational therapists should also assess: mental, sociocultural and spiritual components, occupational roles and expectations, environmental demands, and occupational performance.

With respect to occupational performance assessment, two instruments have been developed within the past five years. These are the Canadian Occupational Performance Measure (COPM) (Law et al., 1991) and the Occupational Therapy Functional Assessment Compilation Tool (OTFACT) (Smith, 1990) in the States. The COPM is the measure investigated in this dissertation.

The main difference between the COPM and OTFACT is in the approach taken. The OTFACT is more quantitative while the COPM is more qualitative. The OTFACT uses computer-based scoring and hierarchical decision nodes for setting performance criteria while the COPM uses a client-centred, self-report approach to measure the performance of activities related to the daily life of clients. The COPM, since its publication in 1991, has been well received by occupational therapists in Canada, as well as in the United States (Law, Polatajko, Pollock, McColl, Carswell, & Baptiste, 1994; Pollock, 1993; Pollock et al., 1990).

The COPM, built on the constructs of a client-centred approach and occupational performance (Law, Baptiste, McColl, Opzoomer, Polatajko, & Pollock, 1990; Law et al., 1991), was constructed for clinical assessment and treatment and program evaluation, and to provide an outcome measure in occupational therapy. Despite its significance in the measurement of clinical outcomes and to the profession as a whole, evidence gathered on the COPM's psychometric properties has been limited (Pollock, 1993; Law et al., 1994). However, given its recency, this is hardly surprising. The need for further research in this area, then, is obvious. The validation activities conducted in the present research are thought to capture the spirit of scientific evaluation of the assessment instrument, theory building, and professional integrity.

Many authors have stressed the importance of using valid, reliable, quantified, and scientifically-based assessment instruments (Campbell, 1989; Christiansen, 1991b; Ottenbacher, 1987). With reference to the COPM, the establishment of evidence of construct validity, reliability, and clinical utility is a basic necessity of the instrumentation processes. Such vigorous evaluation of the use and interpretation of the COPM will contribute to the useful, truthful, and meaningful use of the assessment results which ultimately will benefit the clients who consume the services provided by occupational therapists. Besides these contributions to the instrument and benefits to the clients, findings of this study will enrich the knowledge base of the Canadian model of occupational performance. The substantive and structural components of construct validity will shed some light on the theoretical framework and its application to clinical assessments. Furthermore, the validation model adopted in this study is a novel attempt to apply Messick's (1993) concept of validity in educational and psychological disciplines to the practice of occupational therapy and rehabilitation. The application of Messick's model will certainly inject new perspectives on clinical instrumentation and validation in the assessment and treatment of human performance deficits.

Limitations of the Study

Recognizing that validity and validation processes are "evolving" and "continuing" (Messick, 1993, p.13), generalization of the results of this study was limited by the sampling method, research design, and the resources that were available to the researcher. First of all, evidence of construct validity of the COPM was limited by the participants of this study. The participants were cognitively competent adult clients with physically disabilities, particularly those with orthopedic and stroke problems. Although the decision to select the orthopedic and stroke groups purposefully maximized the differences in the levels of disabilities and treatment modalities of these two groups so that better generalization of results could be made, appropriateness of this restriction needs further investigation and research.

Second, the non-probability sampling method used and the small number of clients involved in this study have limited the generalization of the evidence gathered to other orthopedic and stroke clients. Readers should be cautious and ensure that there are similarities between characteristics, such as age, types of setting and disability, and life roles, of the clients who participated in this study and their own target groups.

Third, the statistical conclusions reached in this study were based on limited sample sizes, particularly in the stroke group. Although preliminary analyses of data showed that the differences in characteristics between clients in the two groups were insignificant, readers should be aware that statistical inferences made in this report were based on small effect sizes.

Fourth, the themes identified in the qualitative analyses revealed only the subjective feelings and evaluations of the participants in this study. No generalization of the results was attempted even though some of the findings were triangulated with their quantitative counterparts. Readers should be selective in extracting evidence.

Finally, the content and strategies of this study were confined by Messick's model of validation. In addition, only those components of Messick's model which were fundamental to clinical instrumentation were employed. Consequently, the evidence gathered is by no means exhaustive. Rather, the researcher regards it more as an impetus to stimulate more research in substantiating the use and interpretation of the COPM.

Organization of the Chapters

Altogether, there are seven chapters in this dissertation. Chapter II is the literature review which provides a fundamental description of the theoretical model of occupational performance and the content and scoring system of the Canadian Occupational Performance Measure. In addition, a thorough discussion is given on Messick's model of validation and its potential application to the evaluation of clinical assessment instruments in the medical and rehabilitative fields.

Chapter III gives detailed descriptions of the methods of investigation used in this study. Different strategies employed to gather different sources of evidence are discussed as well as justification of their use wherever appropriate. This chapter ends with a review of the characteristics of various instruments used in the data collection.

Psychometric properties of the three clinical instruments selected as the "criterion" for establishing convergent and discriminant validity with the COPM are discussed.

Chapters IV, V and VI consist of the presentation and analyses of evidence gathered in this study. Chapter IV describes the characteristics of the clients who participated in this study and their results on the COPM and the three criterion instruments. Chapter V consists of the main bulk of evidence of construct validity of the COPM. In view of the massive amount of information to be disseminated, this chapter is further divided into four sections. These are: content-related evidence, substantive-related evidence, structural-related evidence, and criterion-related evidence. Chapter VI presents the results of the utility and procedural review conducted as part of the study.

Chapter VII is the final chapter of this dissertation and provides a detailed account of the suggestions for changes to the existing COPM based on the evidence gathered in the validation process. Methods to achieve further standardization of the testing protocol and quantitative manipulation of the COPM assessment scores are discussed, and conclusions follow.

CHAPTER II

LITERATURE REVIEW

Introduction

This chapter begins with a description of the model of occupational performance on which the Canadian Occupational Performance Measure (COPM) is based. This model includes various factors thought to influence individuals' occupational performance. The second section is devoted to a detailed description of the COPM, including the history of its development, the content of the instrument, administrative procedures, scoring methods, and suggested use and interpretation of scores. The remainder of the chapter reviews various validation models and processes from the classical perspective to the more recent developments in the measurement field. The validation strategies commonly used in health care instrumentation and research are evaluated. The chapter concludes by introducing the procedures and strengths of the strategies which were used to gather evidence of construct validity of the COPM in this study.

The Model of Occupational Performance

General Background

The concepts of occupations and activities are the domains of concern for the practice of occupational therapy. However, the constructs that underlie these concepts have not been systematically studied, and the field seems to be guided more by individual practitioners' preferences or beliefs than by a strong theoretical orientation. It was not until the early 80's that the field saw the emergence of a theoretical framework in which the various domains of concern were interwoven. This shift initially began with the work of Reed and Sanderson (1980). Following this trend, many theorists in the field made significant advances in building different theoretical frameworks (Christiansen, 1991a; Kielhofner, 1983, 1985, 1992; Mosey, 1986; Rogers, 1983; Townsend, Brintnell, & Staisey, 1990). Different theoretical and practice models were built. For example, in the United States Kielhofner developed a model of 'human occupation'. Townsend et al., in Canada, developed a model of 'occupational performance'.

Although the various theorists identified above used different terminologies to describe their work, the term "occupation" is generally agreed as "engagement in activities, tasks, and roles for the purpose of productive pursuit (such as work and education), maintaining oneself in the environment, and for the purposes of relaxation, entertainment, creativity, and celebration" (Christiansen, 1991a, p.26). In other words, all the goal-directed activities that one performs related to daily living are occupational in nature.

In Canada, an overall model of occupational performance has been developed which has been widely adopted by practising occupational therapists (CAOT, 1991; McColl & Pranger, 1994; Reed & Sanderson, 1980). Guidelines for clinical practice have been developed based on this theoretical model. The model holds a contextual and client-centred approach to occupational therapy (Dunn, 1993; Townsend et al., 1990). It presupposes a holistic view of human behavior and stresses the worth of the individual.

In this model of occupational performance, the components of performance are separated into two levels which guides the clinical practice of therapists. The higher level refers to the individuals' day-to-day activities such as self-care, productivity, and leisure (CAOT, 1991). The lower level performance components refer to the building blocks of human functions which include mental, physical, sociocultural, and spiritual abilities. According to one description of the model (CAOT, 1991), these two levels are integrated with one another and with the environment in which the individuals function:

> Occupational therapists practise within a model which accepts the premise that man has a need to be engaged. His engagement takes many forms and roles, each having a crucial effect on his quality of life. The culmination of man's experiences in this conceptual model is the formation of the individual's mental, physical, sociocultural and spiritual self. The essence of a healthy, functioning person is the balanced integration of these four performance components to provide a sense of well-being. Man integrates these components by engaging with social, cultural, and physical aspects of the environment which he affects and by which he is affected. (p.16-17)

As a consequence, individuals' performances in their "occupation" are affected and predicted by their competence in the performance components. Definitions of the three areas of occupational performance (self-care, productivity, and leisure) and four performance components (physical, mental, sociocultural, and spiritual) used in the model are presented in Appendix I.

The literature on the model of occupational performance is mostly descriptive in nature. It does not provide a detailed account of the structure of the model. Relationships and interactions between and within occupational performance and performance components are not clearly analyzed and explained (McColl & Pranger, 1994). In order to provide the reader a better understanding of the model, the recent literature on human performance and performance deficits is reviewed and summarized in the next few sections.

Core Values and Assumptions of the Model

The philosophical values and assumptions of occupational therapy have had a pervasive influence on the development of the occupational performance model (Mosey, 1986). This section will identify and explain the values and assumptions upon which the model is built, leading to a further identification of the underlying construct of the COPM. This in turn guides the process of gathering the evidence on the content-related and substantive-related validity of the instrument.

Mosey (1986), in her book Psychosocial Components of Occupational Therapy, summarizes the values and assumptions that have existed in the field. Basically, these values and assumptions can be divided into components that refer to the individual and those that relate more generally to occupational therapy as a clinical intervention. One assumption is that individuals actively interact with their human and non-human environments. There are two prerequisites that enable individuals to interact with their environment. One is an individual's right to a meaningful existence, and second is an individual's ability to make choices within the context of certain social constraints. From this perspective, an individual can only be understood within the context of his/her family, community, and cultural groups. In addition, this perspective stresses an individual's inherent needs for work, play, and rest through which their potentials and satisfaction can be reached. At an internal level, what an individual can accomplish is influenced by his/her stage-specific maturation process, social nature, and cognitive structures. Individuals are regarded as dysfunctional when they are not able to reach their goals and fulfil their potentials.

According to Mosey (1986), occupational therapy is the intervention which promotes people's functional independence thus enabling them to come closer to realizing their goals. Occupational therapy as a clinical intervention is directed toward facilitating an individual's occupational performances (major social roles and functions) and developing performance components (abilities and skills) that are fundamental to those performances. The interaction between the individual and occupational components can be seen in the extent to which any kind of intervention is dependent upon the needs of a particular client at any given point in time.

The values and assumptions described above are reflected either directly or indirectly in the Occupational Therapy Guidelines for Client-centred Practice (CAOT, 1991), which is the official document on the Canadian model of occupational performance. McColl and Pranger (1994) summarize the two main values and assumptions inherent in this model: the basic values are that 1) there is an "inherent worth of the individual" and 2) an individual is viewed as "a whole person, whose parts cannot be considered in isolation of one another" (p.16). In a similar perspective, the two theoretical assumptions are that 1) people "have a need to be engaged in activity" (p.16) and 2) an individual is in "a dynamic...constantly changing system" (p.17). These values and assumptions of occupational performance are the basis upon which the theoretical model and model of professional practice are developed.

In the Test Manual (Law et al., 1991) and other literature (Law et al, 1990, 1994; McColl & Pranger, 1994; Pollock, 1993), it is noted that the COPM was built on the model of occupational performance. In other words, the values and assumptions described above guided the collection of evidence on construct validity of the COPM. So, as a clinical instrument measuring clients' occupational performance, the domain of the COPM's test construct should be consistent with the model of occupational performance. To gain a more in-depth understanding of this construct, various factors which are known to influence individuals' occupational performance are discussed next.

Factors Influencing Occupational Performance

The model of occupational performance (CAOT, 1991) suggests a two-level construct of human performance. Occupational performance is interpreted as the outcome of the integration and interaction of four performance components, namely physical, mental, sociocultural, and spiritual. The notion that the individuals' performance components are attributes of their occupational performance prompts the author to explore the effects of each component, or factor upon individuals' functioning, and the nature of interactions among these factors.

Despite a number of research publications in recent years, the literature on the theoretical construct of the Canadian model of occupational performance is still limited. Among the publications to date, the person-environment-performance framework of human performance deficits proposed by Christiansen and others (Christiansen & Baum, 1991) resembles the assumptions and concepts underpinning the model of occupational performance.

The person-environment-performance (PEP) framework developed by Christiansen (1991a) gives a comprehensive depiction of how various factors of individuals influence their occupational performance. According to the framework, factors are categorized into the characteristics of individuals (person), the unique environments in which they function (environment), and the nature and meaning of the activities, tasks, and roles that they perform (performance). Christiansen's framework in fact is a collection of concepts originating from various theorists in the field such as

Clark (1979a, 1979b), Howe and Briggs (1982), Kielhofner and Burke (1980), and Reed (1984).

Personal Factors Influencing Performance

Characteristics of a person can be divided into two categories. The higher level factors which affect an individual's ability to control their performance are: sense of competence, locus of control, and satisfaction. The lower level factors, called "intrinsic enablers of performance", include psychological and cognitive, sensory and perceptual, neuromotor, and physiological factors.

Sense of Competence. An individual's sense of competence is thought to impact most on performance. Competency is based on the premise that a human being has an innate drive toward mastering the environment. Individuals are motivated or driven to act and interact with their environment through occupation. Sense of competence is developed when individuals experience successes in dealing with environmental challenges. Self efficacy and personal effectiveness are sources of personal satisfaction which result in competent performance within the environment.

Locus of Control. The behavior of individuals is influenced by their expectation of successful outcomes, perception of the importance of the outcome, and locus of control. Internal locus of control prompts people to act on the basis of their satisfaction or their own incentives that lie within their actions (Mosey, 1986). External locus of control is the result of experience which originates from the environment and which may or may not be controlled by the individuals and relate to their actions.

Satisfaction. Closely related to competence and locus of control, is an individual's satisfaction. Satisfaction is described as the outcome of the situation in which deprived needs are gratified (Mosey, 1986). Maslow's (1970) hierarchy of needs provides the links between inherent needs of people and their satisfaction and gratification. However, it is argued that Maslow's thesis is too narrow because it concentrates too much on the intrinsic factors of individuals without the acknowledgement of extrinsic effects emerging from the environment. In contrast to Maslow, Fuhrer, Rintala, Hart, Clearman, and Young (1992) suggest that people feel more satisfied when they feel their life is under more of their own control. In their study, Fuhrer et al. reported significant moderate correlations between life satisfaction and perceived control ($r=.50$, $p<.01$), self-assessed health status ($r=.41$, $p<.01$), and social support ($r=.32$, $p<.01$). In a study conducted by Yerxa, Burnett-Beaulieu, Stocking and Azen (1988), results showed that self-reported independence is highly correlated with an individual's satisfaction with life (ranged from $r=.75$, $p<.001$ for social and recreation skills to $r=.45$, $p<.05$ for self-care activities). In addition and similarly, a significant positive relationship was found between satisfaction with performance and overall life satisfaction (Yerxa & Baum, 1986). In summary, an individual's satisfaction in general or satisfaction specifically with performance is closely related to their level of independence and the level of control that they perceive they have over their own actions.

Intrinsic Enablers of Performance. Intrinsic enablers include abilities and skills of individuals such as physiological, sensory, neuromotor, cognitive and psychological functions. Abilities are either innate or determined genetically; skills are either learned or obtained through practice and experience. Both abilities and skills are unique to each individual and are requisite of one's performance of life tasks. Christiansen's descriptions of the "enablers" are equivalent to the four performance components in the occupational performance model.

Activity, Tasks, and Role Factors Influencing Performance

Christiansen (1991a) defines occupational performance as:

> ...the day-to-day engagement in occupations that organize our lives and meet our needs to maintain ourselves (self maintenance), to be productive (productivity), and to derive enjoyment and satisfaction within our environment (leisure). (p.27)

This definition is more a descriptive and operational representation of occupational performance than the definition stated in the Occupational Therapy Guidelines for Client-centred Practice (CAOT, 1991). However, the domain of concerns addressed by both are largely the same, namely, self-care, productivity, and leisure. According to Christiansen, performance is influenced by the nature and meaning of various activities and tasks, and the role responsibilities of individuals.

Activities and Tasks. Activities and tasks are not meaningful if they are isolated from the roles and social context of individuals. This argument is based on the premise that all goal-oriented activities related to daily living are associated with certain performance expectations (Christiansen, 1991a). These performance expectations are determined by internalized values which accompany individuals' social roles (Kielhofner, 1985). In other words, this conceptual framework suggests a relationship between individuals' values, social role, and performance. Christiansen described this relationship in an occupational performance hierarchy. He placed social roles of individuals at the highest level of performance. Roles are further suggested to be occupied by individuals, defined by their own performance expectations, and viewed as attributes of their performance. For example, in a house keeping or home maintainer role (Oakley et al., 1986) individuals are expected to prepare family meals (performance expectation), and meal preparation is considered to be the role responsibility of the home maintainer. An activity, on the other hand, is the basic unit of occupational performance which consists of specific goal-oriented behavior directed toward an individual's performance. However, activities are not contextual nor specific to any particular roles of individuals. For example, the activity of preparing a meal can be done by all people regardless of their particular life roles. A task is a set of activities with a shared purpose that is recognized by individuals. For example, making salad is a task that contributes to the preparation of a meal. Hence the relationships among tasks, activities, and role responsibilities are demonstrated.

Roles. The concept of roles in the context of occupational therapy is slightly different from the concept of social roles in social psychology. Roles are defined as an organized pattern of behavior that is characteristic and expected of the occupant of a defined position in a social system (Mosey, 1986). Some examples of these defined roles are teacher-student, child-parent, husband-wife, and therapist-client. Theorists of the occupational behavior tradition have argued that roles help to organize productive behavior by providing a personal identity, conveying social expectations for performance, organizing use of time, and placing the individual within the social structure (Rogers, 1983). The social expectations are formed by both society and the role occupant. As a consequence, one's satisfaction with the performance of valued roles is based on internal as well as external appraisals (Christiansen, 1991a).

Besides the pattern of roles, the consistency between an individual's abilities and performance of particular roles is of interest in occupational therapy. Studies have shown that disability resulting from trauma and illness commonly leads to changes in role function, role patterns, and the balance of roles of individuals (Hallett, Zasler, Maurer, & Cash, 1994; Oakley, 1982). These changes range from minor modifications to obliteration of all major life roles which are related to personal, social, and environmental factors (Giles, 1994).

Environmental Factors Influencing Performance

Although the environmental factors are extrinsic to individuals, they tend to exert tremendous influences on an individual's level of performance. Rogers (1983) called these environmental factors "enablers of human performance". Literature on occupational performance has revealed close relationships among person, activities, and environment. Christiansen (1991a) classified the environmental factors into cultural influences, social dimensions, and physical environment. These factors influence an individual's decisions to interact with their surroundings and have an impact on the quality of their performance.

Cultural Influence. There is no single definition that captures the meaning of culture in a precise fashion. Rather, definitions available in the literature provide descriptions of phenomenon under the influence of culture. Culture, as defined by Mosey in 1986, is "a set of shared understandings held in common by members of a group" (p.173). The culture of a group is the sum of its morally forceful understandings acquired through learning and shared by the members of that group. Theoretically speaking, it is difficult to operationalize the construct of culture. Nevertheless, Krefting and Krefting (1991) identified some important factors that are believed to influence the performance of an individual. From the list of factors provided by Krefting and Krefting, nine are relevant to be relevant to the context of the present research. These nine are family structure, economic status and history, educational background, age, marital status and history, vocational status and history, religion or spiritual orientation, immediate environment, and health-related experiences. When the concept of cultural influence is transferred into the practice of occupational therapy, Krefting and Krefting suggest that therapists, when evaluating the performance of a client, must assess the cultural identity of the client and use this information to plan culturally sensitive interventions.

Social Environment. Despite the difficulty of separating the social environment from the cultural and physical components as an independent entity for discussion, Davidson (1991) clearly presents the concepts of social influence in terms of social networks and socially defined activity patterns, beliefs, and expectations.

The term "social network" is a way of conceptualizing dimensions of social environment that represent a specific type of social relation linking a defined set of people (Davidson, 1991). The people in the social network are selected for inclusion because of the attributes that they possess, for example, a member of a family, an ethnic group, a profession, or a neighbourhood. One of the major functions of a social network is to provide social support. At a personal level, individuals within a social network are capable of providing support which enhances social identity, provides emotional aid, material aid, and information needs. However, Wellman (1978) noted that not all of the relationships in a network may be supportive in nature. For instance, some relationships could manifest themselves in power struggles, rumour spreading, and other unsupportive events. Besides physical and emotional components, social support is found to play an important role in the development and maintenance of the individuals' functional performance.

Physical Environment. The impact of the physical environment on the performance of individuals is discussed last in this section because the physical environment is the arena in which the cultural and social effects occur. The physical environment refers to the structural elements of a home or work setting, interior and furnishings, tools and equipment, and environmental hazards. All of these factors, as a collective unit, shape individuals' behavior, social interaction, and activity patterns (Spencer, 1991).

At a concrete level, the physical environment can be an impediment to the mobility and activities of the daily life of individuals due to architectural barriers and/or inappropriate use of tools and equipment. At more abstract level, influences of the physical environment extend to hamper the psychological and emotional well-being of individuals. This could be due to the incongruence between individuals and their environment, such as the atmosphere of their surrounding and environmental press.

Summary on the Factors Influencing Performance

The theoretical framework of the model of occupational performance suggests a multi-factorial structure among the factors which are believed to affect individuals' performance of their day-to-day activities. The structures, as depicted by various authors, have not been well operationalized, quantified, or validated. The structure that does seem to emerge from the literature is a two-level construct. This construct describes the characteristics of individuals with a unique set of basic abilities and learned skills at the lower level (physical and mental performance components), and functional activities relevant to individuals' life roles and expectations in self-care, productivity and leisure at the higher level. Interacting with the physical and social environments in which individuals function, the abilities and skills of individuals are adapted to different expectations, patterns of activity, and life roles (socio-cultural and spiritual performance components).

The result of the complex integration and interaction between "persons" and their "environment" is competence in their occupational performance. When the perceived level of competence has been reached, individuals are likely to express satisfaction with their performance.

In this study, the model of occupational performance was used as the theoretical construct against which the COPM was evaluated. It was used to guide the collection of evidence on the content-related and substantive components of construct validity of the instrument (Messick, 1989). The results of this study help shed some light on the content and structure of the occupational performance model.

Canadian Occupational Performance Measure (COPM)

The COPM (Law et al., 1991) was developed by the Department of National Health and Welfare and the Canadian Association of Occupational Therapists Task Force. It is an individualized criterion referenced assessment tool designed for measuring changes in occupational performance of clients receiving occupational therapy. The term "occupational performance" refers to clients' competence in the performance of activities related to day-to-day achievement of self-care, productivity, and leisure (CAOT, 1991).

The COPM uses a client-centered and self-report format. Therapists administering the instrument play a facilitative role in helping their clients respond to the instrument (Law et al., 1991). Various authors have commented that the COPM is superior to other more traditional assessment tools used in occupational therapy and the rehabilitation field (Christiansen, 1993; Law, 1993; Law et al., 1990; Mathiowetz, 1993). Traditional clinical assessment tools have been criticized for being too performance component oriented, not client-centred, not generic, and not sufficiently holistic in the assessment of clients' role function and environment (Christiansen, Schwartz, & Barnes, 1988; Law et al., 1990). According to authors of the COPM, most of the critical characteristics of a good measure of occupational performance were incorporated into its construction. The motivation behind the present study was to examine the evidence on the validity and reliability of inferences drawn from the scores and information yielded by the COPM. If the interpretations and uses of the COPM in measuring clients' occupational performance are credible, clinicians can have confidence in the COPM as a clinical outcome measure.

Content of COPM

The COPM uses a self-report format and a semi-structured interview method in its assessment process (Law et al., 1991). Five components or steps are assessed: problem definition, problem weighting (Importance rating), scoring (Performance and Satisfaction ratings), re-assessment, and follow-up. Each component or step is described below.

Problem Definition (Step 1). In this step, clients are asked to list the activities that they find difficult to perform in the areas of self-care, productivity, and leisure. In the process, clients are prompted to think of the activities they need to perform or are expected to perform and whether or not they are able to perform them satisfactorily.

Problem Weighting (Step 2). In step two, clients are asked to rate the activities that they identified in Step 1 in terms of importance in their life. Importance of each activity is rated on a ten-point rating scale (from 1 to 10). The most important or pressing five problems (those with highest Importance ratings) are then selected to enter the next step of ratings by the clients. The Importance ratings become the weighting factors of the activities identified.

Scoring (Step 3). Here clients are asked to evaluate their own current performance on the activities that they identified in Step 2 against a ten-point (from 1 to 10) Performance rating scale (Performance 1). Clients are then asked to rate their satisfaction with their current performance of those activities on the ten-point Satisfaction rating scale (Satisfaction 1).

The computations of Performance and Satisfaction subscale scores of each identified activities are as follows:

Performance = Importance X Performance 1

Satisfaction = Importance X Satisfaction 1

The item score for Performance and Satisfaction of each activity identified for rating can range from 1 to 100. The same scoring process is repeated for all the identified activities on the COPM (a maximum of 5). The total Performance Score 1, called the Performance subscale score, is the mean of all the item Performance scores. The total Satisfaction Score 1, called the Satisfaction subscale score, is the mean of all the item Satisfaction scores. Both subscale scores have a range of 1 to 100.

Re-assessment (Step 4). The COPM is designed to measure changes in client's occupational performance throughout the process of occupational therapy. A re-assessment is conducted at an appropriate interval following the initial assessment (step 1, 2 and 3) and intervention. The Test Manual of the COPM (Law et al., 1991) does not specify the time interval between the first assessment and re-assessment. Instead, it suggests an eclectic approach which is determined by clients and their therapists.

During re-assessment, clients are asked to rate those activities that they identified during the first assessment. They are asked to rate their current Performance and Satisfaction by using the same rating scales. The scores from this second assessment entitled Performance 2 and Satisfaction 2 are multiplied by the original Importance ratings to yield the second item Performance and Satisfaction scores. The mean of item scores on the two scales forms the total Performance Score 2 and total Satisfaction Score 2.

Change in Performance = Performance Score 2 - Performance Score 1

Change in Satisfaction = Satisfaction Score 2 - Satisfaction Score 1

The changes in occupational performance of a client are measured by the two subscale scores of Change in Performance and Change in Satisfaction.

Follow-Up (Step 5). After the re-assessment, therapists and their clients make plans for either follow-up, treatment continuation, or discharge. According to the Test Manual, follow up assessment repeats the original assessment process starting from Step 1 - Problem Definition. On this occasion, therapists ask their clients whether there are problems in occupational performance that remain unsolved, or whether new problems have emerged over time. In this case, the process continues by using new COPM forms. The Test Manual specifies that the COPM can be repeated with clients as many times as deemed necessary.

In the Test Manual, the authors of the COPM have stressed the importance of interaction between therapists and their clients during all the assessment processes. Therapists are encouraged to integrate the relevant assessment findings into the planning and implementation of clients' treatment programs.

Although the COPM has the potential to become the model client-centered assessment instrument in occupational therapy, its psychometric properties have not been widely studied in the field. By the end of 1994, only a handful of reports had been published on pilot studies examining its psychometric properties. Nevertheless, various authors have collected positive feedback in the use of the instrument (Law, Polatajko, Pollock, McColl, Carswell, & Baptiste, 1994; Pollock, 1993). However, findings from these studies are mainly descriptive in nature. So, in order to implement a more thorough evaluation of the COPM, the author has attempted to adopt a validation process with in-depth strategies (Loevinger, 1957; Messick, 1989; Shepard, 1993) to establish evidence of construct validity on the use and interpretation of the COPM.

COPM and Use of Proxy

According to the Test Manual (Law et al, 1991), the COPM can be used "across all developmental levels" and "disability groups" (p.10). The instrument claims to be appropriate in measuring the occupational performance of a diversified clientele including all age and clinical diagnostic groups. Such a claim inevitably raises questions about how clients classified as cognitively incompetent (e.g. those with head injuries and Alheizmers' clients, and small children) are to be interviewed? The Test Manual offers this recommendation:

> ...you may decide to use the client's family or primary caregivers as the respondent. They may have to answer questions on the client's behalf.

> The respondent should be instructed to consider their own environment and expectations, as well as what they perceive to be the needs of the client. (p.21)

The underlying assumptions of the these recommendations are that 1) the client's family or primary caregivers' "perceptions" of the client's occupational performance is similar to client's own perceptions; and 2) the extrinsic factors (Christiansen & Baum, 1991) affecting the client's performance, such as social and physical environment (areas in which family and caregivers can accurately respond), are more important in determining his/her occupational performance than the client's intrinsic enablers of performance, such as motivation, cognitive functioning and sensori-motor functions (areas which only the client can accurately respond and/or be measured). However, these two assumptions, as the author will argue, seem to be problematic and deviate from the ideology of the "client-centred" approach.

With reference to the first assumption, it is obvious that no two persons have exactly the same perceptions. This is true even for people with close relationships such as spouses and siblings. The second assumption reflects a shift of ideology from an individualized (client-centred) and criterion-referenced basis to a professionally judged and norm-referenced basis. Perceptions and information given by caregivers and family members may work against the truthfulness of the assessment, such as a tendency to report lower Performance and Satisfaction ratings if clients are perceived to be a burden. At the other extreme, caregivers and family members may have a tendency to report inflated ratings on Performance and Satisfaction as a reflection of social desirability or if they have implicit expectations for the clients to undergo a faster recovery process. Review of the literature clearly suggests complicated and dependent relations among all these factors. However, these issues were neither addressed by the model of occupational performance nor by the authors of the COPM (CAOT, 1991; Christiansen & Baum, 1991; Law et al, 1991).

Although the COPM recommended that therapists use proxies in situations where clients are unable to cope with the COPM assessment process, the present validation study was restricted to clients who were able to participate in the testing process without any assistance from a proxy. This eliminated the confounding effects of using a proxy as described above. As a result, the findings of this study can not be generalized to the situation in which a person other than the respondent is interviewed.

COPM and Cognitive Competence

The term "cognitive competence" is an arbitrary concept which is set against certain test behaviors or performance tasks. It is obvious that an individual requires different cognitive abilities to differentiate between two colors or to decide which of two activities are more important to enhance their quality of life. But how different are the cognitive abilities required for individuals to solve these two types of problems? Newell and Simon's (1972) theory on problem solving seems to provide a good foundation source from which to answer these questions.

Newell and Simon postulate that there are several basic elements in the ways individuals process information. These elements include storage time, retrieval time, and the capacity of short-term and long-term memory. According to Newell and Simon, short term memory (STM) places a constraint on the number of sequential operations that can be carried out because of the limited capacity in any individual's STM. More complicated tasks involve more sequential operations. Hence individuals require a larger capacity in their STM in order to successfully solve these more complicated problems. So, using the example above, the tasks of deciding which activities are more important in enhancing one's quality of life requires higher cognitive functions than does differentiating between colors.

A complicated task, such as "be able to cook for oneself", requires a client to consider multidimensional attributes. These are:

1) what activities are required (role expectations)?
2) what is preferable: dependence or independence?
3) what are the consequences if those activities are not performed adequately?
4) what resources and help are available (support system)?
5) what does "importance" mean?
6) what number should I give to represent the degree of Importance for those activities?

This list is far from exhaustive, but it illustrates the complexity of some of the operations that may occur in STM when individuals are asked to perform a similar task as in Step 1 - Problem Definition in the COPM.

Newell and Simon's model of information processing also makes inferences about the roles of long term memory (LTM) in the performance of problem solving tasks. Unlike STM, LTM does not have a limited capacity. However, there is a time constraint on registration and retrieval of information between the STM and LTM. The longer the time required to process the information in a task, the greater the difficulty in remembering the steps for solving that problem. A person's competence in solving the problem diminishes with the greater the amount of time required to process the task or information.

Besides STM and LTM, an individual's competence in performing problem solving tasks requires other cognitive abilities. For example, Reed (1992) notes fluency in generating possibilities, retrieval of solution patterns, and knowledge of principles as important cognitive abilities in solving problems. Sternberg (1986) specifies the importance of encoding, inference, mapping, and application. Still other cognitive abilities include orientation, comprehension, attention and concentration, verbal and abstract concept formation, and the use of receptive and expressive language skills (Sattler, 1992).

In the COPM, the testing process requires the clients to respond in a semi-structured interview. Clients respond to tasks in which they must identify the activities that they find difficult to perform, rate the importance of these activities, rate their perceived performance in the activities, and rate their satisfaction with their performance. These tasks all demand the clients' competence in decision making and problem solving. The clients are expected to have good knowledge of their role functions and expectations, be able to fully assess their performance, and have good insight into their own feelings. Moreover, they are assumed to be capable of integrating all these facts and perceptions, and making multi-modal decisions about their occupational performance.

In view of the high cognitive functions that are required to complete the assessment processes in the COPM, clients who participated in this study were screened by the Neurobehavioral Cognitive Status Examination (NCSE) (Northern California Neurobehavioral Group, 1988) to eliminate those with identifiable cognitive impairment. The details on the administration and psychometric properties of the NCSE are discussed in Chapter III. Because the sample of this study may not represent the target population of the clients specified by the COPM Test Manual, the generalizability of the results and inferences generated by this study are inevitably impeded.

Validation Process in Occupational Therapy

Despite the fact that the use of standardized and valid assessment tools is considered to be important in the professional practice of occupational therapy (Alberta Association of Registered Occupational Therapists, 1990; Benson & Clark, 1982; Dunn, 1989; Law, 1987), therapists have not made much progress in constructing and validating the instruments on which they must rely (Campbell, 1989). However, among occupational therapists, there has been a general increase in the awareness of the necessity of establishing acceptable levels of reliability and validity of the inferences drawn from the scores yielded by the instrument (Smith, 1992).

Literature in the area of measurement in occupational therapy shows two methodological weaknesses that are common to the majority of the validation studies reported. First, many authors advocate a "classical" model of validity (American Psychological Association, 1954) which assumes a fragmented concept of "validity" (for example: Deusen, Shalik, & Harlowe, 1990; Exner, 1993; Katz, Itzkovich, Averbuch, & Elazar, 1989; Mathiowetz, 1993; Rothman, Hedrick, & Inui, 1989; Wood-Dauphinee, Opzoomer, Williams, Marchand, & Spitzer, 1988). Many measurement experts and theorists have demonstrated the major problems encountered when dividing validity into content, criterion, and construct components (Crocker & Algina, 1986; Cronbach, 1988; Messick, 1989; Shepard, 1993). First, there is an inherent and substantial overlap among the different types of validity. Second, the different components are not independent of one another. According to more recent thought, a unified concept of validity (APA, 1985; Cronbach, 1984, 1988; Messick, 1980, 1989) is a more comprehensive and logical replacement for the traditional "trinity" concept. The unified concept does not merely impose the centrality of construct validity as the ultimate goal of validation to which evidence from various sources is gathered. Rather, it secures a basic presupposition of test validation as a research process involving the meaning, theoretical underpinnings and relationships, interpretations, inferences, and implications for action of test scores in use. One does not validate an instrument once and for all. Rather, one investigates the validity within the theoretical context of interpretation and use.

The second methodological weakness is that exclusive emphasis has been put on relying on "statistical" methods to gather evidence of validity, especially in establishing evidence of "criterion validity" (for example: Edwards, 1990; Evans & Salim, 1992; Fillenbaum & Smyer, 1981; Houston, Williams, Bloomer, & Mann, 1989; Primavera, Novello, Finocchi, Canevari, & Corsello, 1990; Tombaugh & McIntyre, 1992). Under the notion of the "statistical" concept of validity, validity coefficients reported are predominantly criterion and predictive in nature. The coefficients, computed by correlational procedures, would be different when data collected for establishing evidence of validity are gathered under different conditions. Factors external to the test and tested individuals, such as differences in clinical settings, time of testing, and slight changes in administrative procedures, are all known to produce different validity coefficients (Gulliksen, 1950; Meehl & Rosen, 1956). In other words, the validity of a test does not merely tap "true scores" or traits of individuals. Instead validity is influenced by other factors extraneous to the test that were never intended or claimed by the test developer. Moreover, the computation of correlation coefficients to describe relationships between test constructs is too superficial. Two sets of unrelated data (representing two unrelated constructs) could, in fact, correlate highly with one another. For example, cognitive functioning and height of children are known to be independent of each other. However, in a normal developmental process, cognition and height of children are found to be positively correlated since they both increase naturally with age.

Loevinger, as early as 1957, stressed the importance of gathering trait-oriented evidence and using more in-depth approaches to investigate the validity of tests. In recent years, authors in measurement have reiterated the notion that empirical relationships are necessary, but not sufficient, to establish validity of a test score inference (Cronbach, 1988; Shepard, 1993). More vigorous strategies such as factor analysis (Loevinger, 1957; Messick, 1989), multitrait-multimethod analysis (Campbell & Fiske, 1959), path analysis and structural equation modeling (Puderbaugh & Fisher, 1992), expert judgement by panel review (Betz & Weiss, 1976; Carswell, Carson, Walop, & Zgola, 1992; Ebel, 1967; Thorn & Deitz, 1989), protocol analysis of test process (Ericsson & Simon, 1985; Meyers, Lytle, Palladino, Devenpeck, & Green, 1990), and meta-analysis (Cohen, 1977; Ottenbacher & Barrett, 1990; Ottenbacher &

Tomchek, 1993) are recognised as powerful strategies to establish "content-related" and "criterion-related" evidence of construct validity.

In occupational therapy, the fragmented "classical" model of validity and the "statistical" method are still the "gold" standards guiding practitioners in carrying out test validation. The author argues that the "myths" of the notion of validity in occupational therapy may be attributed to the lack of a thorough understanding of the underlying concept, and the lack of rigorous strategies by the field for establishing evidence of construct validity of test use and interpretation. In view of these problems, it is the purpose of this study to apply Messick's (1989) model of validation, an in-depth and comprehensive approach, to the evaluation of the COPM in occupational therapy.

Messick's Concept of Validity - Toward a Validation Model

Messick's (1989) unified concept of validity has emerged from an educational and psychological measurement context. It has not been widely applied to evaluate the interpretation and use of clinical assessment tools in occupational therapy. The application of Messick's model of validation to the COPM is thought to be appropriate for two reasons. Firstly, as with other clinically-based disciplines, occupational therapists assess clients and interpret these assessment results for well defined purposes. Secondly, the assessment results on which occupational therapists make their decisions have an impact upon clients, their family, and society in general.

Messick (1989) defines validity as "the degree to which empirical evidence and theoretical rationales support the adequacy and appropriateness of interpretations and actions based on the test scores" (p.13). His definition reflects a few underlying presuppositions on the notion of validity. Messick makes these claims about validity:

1) Validity is a scientific concept, and validation uses the natural science approach. The concepts of research design, hypothesis testing, and theory testing are applicable to the process of validation.
2) Validity is a unitary concept which needs the convergence of evidence both from empirical and theoretical sources.
3) Validity exists along a continuum, and should be described in terms of degree rather than all or nothing. Hence validity has an evolving property and validation is a continuing process.
4) Validity is associated with the interpretation and use of the test score but not the test itself.
5) Validity should be extended beyond test score meaning to include relevance, utility, value implications, and social consequences.

Messick proposes two dimensions of validity that interact with each other as shown in Figure 2.1. The first dimension is the source of justification for testing. It is divided into appraisal of evidence ("evidential basis") and of consequences ("consequential basis"). The "evidential basis" refers to the collection of validation evidence by various techniques such as correlational and experimental methods. The "consequential basis" addresses the issue of values and the implications associated with the test, testing process, and test results. The second dimension encompasses the function and outcome of testing: test interpretation and test use. Test interpretation refers to the purpose of using test results to describe a person, such as an individual's independence level and performance. Test use, on the other hand, refers to decision making based on an interpretation of test results, for example, to assign a client to a particular remedial program or discharge a client from an institution.

Figure 2.1

Messick's Facets of Unified Validity

	TEST INTERPRETATION	TEST USE
EVIDENTIAL BASIS	Construct Validity	Construct Validity + Relevance / Utility
CONSEQUENTIAL BASIS	Value Implication	Social Consequences

Note. From Messick (1989).

Messick (1989) does not abandon the concepts of "content", "criterion", "structural", and "substantive" validity in his unified model. He has incorporated the ideas of prominent writers in measurement such as Cronbach (1971) and Loevinger

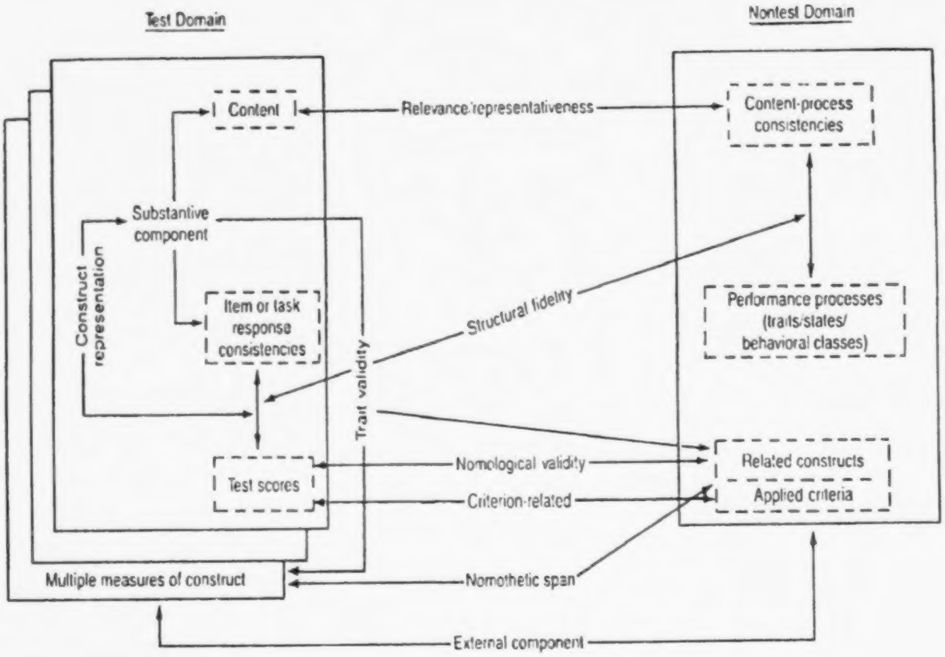
(1957) in his reformed validation model and has integrated the aforementioned concepts. Under the new model, various classes are regarded as sources and components of evidence which are "subsumed under the rubric of construct-related evidence" (p.20). The term "construct validity" retains its original meaning from the classical model, but also includes the evidence and rationale supporting the trustworthiness of meaning of scores in terms of explanatory concepts that account for both test performance and relationships with other variables.

As indicated in Figure 2.1, the evidential basis of test interpretation is construct validity. The evidential basis of test use is also construct validity, but it is specified further to include evidence for the relevance of the test to its particular purpose, and for the utility of the test in the applied setting. The consequential basis of test interpretation is the appraisal of the value implications of the construct label, of the theory underlying test interpretation, and of the theoretical implications in which the theory is embedded. Finally, the consequential basis of test use is the appraisal of both potential and actual social consequences of the testing process.

Basically, Messick's model of construct validity is comprised of various sources of evidence including content-related, substantive, structural, and external components (Messick, 1989). Figure 2.2 shows a diagramatic representation of Messick's model of validity. The first three components can be grouped as internal to the test construct. They reflect all aspects of theory such as the expected interrelations among dimensions of the construct and the processes believed to underlie test performance (Shepard, 1993). The external component establishes evidence of the relation of the intended construct to other constructs.

Figure 2.2

Messick's Components of Construct Validity


Test Domain
Nontest Domain
Content
Relevance/representativeness
Content-process consistencies
Substantive component
Construct representation
Item or task response consistencies
Structural fidelity
Performance processes (traits/states/ behavioral classes)
Trait validity
Test scores
Nomological validity
Related constructs
Criterion-related
Applied criteria
Multiple measures of construct
Nomothetic span
External component

Note. From "Validity." by S. Messick. in Educational Measurement (3rd ed., p.50) edited by R.L. Linn. 1989. New York: American Council on Education and Macmillan. Copyright 1989 by the American Council on Education and Macmillan. Reprinted by permission.

Content-related Evidence of Construct Validity

Content-related evidence involves the specification of the nature and boundaries of the construct domain as well as the appraisal of relevance and representativeness of the test items with respect to the construct domain. The prerequisite of this source of evidence is to have a well-defined theoretical construct base on which the instrument to be validated is built. The non-test domain sets the boundaries for the expected behavior tested by the instrument. The behavior constituting the test domain should be sampled from the universe of behavior constituting the non-test domain. For example, the construct of occupational performance sets the parameter and content for the COPM to test clients' performance in the areas of self-care, productivity, and leisure activities (CAOT, 1991). The relevance and representativeness of the test domain, either test items or test behavior, can then be judged in relation to the non-test domain representing the theoretical construct.

According to Messick, the appraisal should take into account the testing procedures that significantly affect test performance including underlying processes, administrative conditions, and criteria for item scoring. A common strategy used in collecting evidence is to apply a consensual expert judgement approach in which the relevance and representativeness of the content and method of assessment are evaluated. In this study, a panel review of practising therapists was the method used to collect evidence of content-related validity. The method of panel review will be further discussed in a later section of this chapter.

Substantive Evidence of Construct Validity

Messick uses Loevinger's (1957) definition in conceptualizing the substantive component of construct validity as "the extent to which the content of the items included in (and excluded from) the test can be accounted for in terms of the trait believed to be measured and the context of measurement" (p.97). Besides content relevance, the inclusion of particular test behavior in a test is judged by its "empirical response consistency". In other words, different items or test behavior in an instrument should tap the behavior or performance under a unitary construct, for example, occupational performance not perceptual or cognitive skills of clients. Henceforth, responses of the clients being tested will be consistent in items or test behavior.

There are several different techniques used to quantify the substantive component of construct validity. These are: 1) the convergent-discriminant strategy which uses tests of similar and dissimilar constructs and formats to correlate with the intended instrument (Messick, 1989); 2) protocol analysis which investigates the underlying mental processing of clients to justify the unobservable test behaviors purported to underlie the test construct (Ericsson & Simon, 1980; Meyers, Lytle, Palladino, Devenpeck, & Green, 1990); 3) other explorative and quantitative methods which include factor analysis, structural equation modeling, and path analysis (Messick, 1989). In this study, protocol analysis and convergent-discriminant methods were used to gather evidence of substantive validity of the COPM. Details of these two methods will be further elaborated in later sections of this chapter.

Structural Evidence of Construct Validity

Structural validity refers to evidence gathered on the internal structure of a test and is divided into two facets: the fidelity of the scoring model to the test construct and the degree of inter-item structure (Messick, 1989). Structural fidelity, using Loevinger's (1957) term, is "the extent to which structural relations between test items parallel the structural relations of other manifestations of the trait being measured" (p.97). An example to illustrate structural fidelity is the Klein-Bell ADL Scale (Klein & Bell, 1979). This scale consists of several sub-scales for each area of self-care functions such as dressing and feeding. Structural fidelity applies in determining the appropriateness of either interpreting the sub-scales as such, or of looking at the composite scores by aggregating the sub-scale scores. Inter-item structure refers to the relationships among various items or test behaviors in terms of the scoring system, assignment of item weight, and the computation of the final scores. Evidence is gathered on whether the item scores, sub-scale scores, and final scores of an instrument reflect the test construct. Other issues such as the method of discriminating among subjects by either norm-referenced or criterion-referenced standards, and the use of cumulative versus class models are also relevant to inter-item structure. Various methods were used in this study to establish evidence of structural validity including quantitative item analysis and panel review.

External Evidence of Construct Validity

The external component of construct validity requires evidence of how the test construct relates to both similar and dissimilar constructs external to the instrument. This component is further divided into two parts, namely trait validity and nomological validity. Trait validity, according to Messick, deals with the fit between measurement operations of a test and the conceptual formulation of the test construct. The multi-trait multi-method technique (Campbell & Fiske, 1967) is the method commonly used to investigate these relationships. The assumption of trait validity is that the measurement of a test construct should not be tied to a particular method of tapping test behavior. That is, the test behavior obtained through different formats, such as direct observation and a self-rating scale, should yield the same interpretations for both test performance and test scores. As a consequence, the correlation among measures of the same construct using different test formats should be high, and should be higher than the correlation of different constructs using the same format and different constructs using different formats.

Nomological validity is based on the notion that the theory of a test construct should prescribe the predicted links between the test scores and measures of other constructs. In other words, the basis of nomological validity assumes that "no single test is a pure exemplar of the construct but contains variance due to other constructs and method contaminants" (Messick, 1989, p.48). Similar to trait validity, nomological validity is commonly illustrated by the multi-trait multi-method technique.

Different from trait validity, the evidence of nomological validity is derived from the correlations between the test under validation and some other tests. Such a quantitative approach enables the differentiation of "shared" variance from "unshared" variance based on which aspects of construct meaning derived from the shared and unshared parts can be explained. In order to simplify the terminology used in this dissertation, the trait and nomological validity are regarded as "criterion-related validity". Correlational methods using the concepts of multi-trait multi-method will be used to investigate the criterion-related validity of the COPM in this study.

Summary

Messick's unified concept of validity is new to occupational therapists. However, in the field of psychology and education, the change from the tri-partite model to the present unified model has been widely debated by theorists and practitioners. Despite the fact that there are still contentious issues about Messick's model (Shepard, 1993), such as whether different facets of validity should be identified, and whether the so-called social consequences of test use should be included as part of validity, the integrated concept of validity is thought to benefit the theory development and practice of occupational therapists. Messick's model gives a firm theoretical background for validity and validation which researchers and practitioners in occupational therapy can use to build their own models of test construction and validation. Based on Messick's model of validation, strategies used for collecting and analyzing the data in this study were panel review (Thorn & Deitz, 1989), protocol analysis (Anderson & Bower, 1973; Simon, 1979), and the multi-trait multi-method (Campbell & Fiske, 1959).

Strategies of Establishing Evidence of Construct Validity

Researchers and theorists in occupational therapy have stressed the need to use new evaluation methods to validate assessment tools, such as Rasch analysis (Fisher, 1993; Law, 1993), factor analysis (Deusen, Shalik, & Harlowe, 1990; Wood-Dauphinee, Opzoomer, Williams, Marchand, & Spitzer, 1988), vigorous analysis on theory-driven assessments (Dunn, 1993; Fisher & Short-DeGraff, 1993), multi-dimensional scaling (Ottenbacher & Cusick, 1993), and the employment of expert panel review (Carswell, Carson, Walop, & Zgola, 1992; Thorn & Deitz, 1989). The effort and needs should be coupled with a better understanding of various common techniques that are used to gather evidence of construct validity. Although many of the techniques require some knowledge in measurement and statistics, the information presented here for the panel review, protocol analysis and multi-trait multi-method matrix will be at a conceptual level.

Expert Judgement by Panel Review

Panel review is a common method that is used by researchers to gather content-related evidence of construct validity. It is one of the best known strategies that occupational therapists have applied to the construction and validation of clinical assessment instruments. For example, Exner (1993), in constructing the In-Hand Manipulation Test (IMT), set up two expert panels composed of occupational therapists to assess the "content validity" of the test. She reported that the test's table of specifications and other information about the test were initially revised based on the feedback from the first panel (of 12 occupational therapists). The second panel's members (consisting of 24 therapists) were asked to review the content of the test by watching videotapes which showed the actual testing process.

The prerequisite of a panel review is a clear test domain and a table of specifications for the instrument of validation. Judges in the panel are asked to evaluate the instrument against pre-determined criteria. The review criteria are not limited to the relevance and representativeness of objectives in the table of specification, but they also related to testing procedures, administrative conditions, and criteria for item scoring (Messick, 1989). Raters or reviewers often use the criteria in the form of a rating scale. The rating scale can be in a nominal format, such as +1.00 for congruence and -1.00 for incongruence, or in an ordinal format, with irrelevant (1) to relevant (5). Content-related evidence is usually derived from the degree of agreement among the raters or reviewers. Many indexes of agreement - percentage of agreement (McDermott & Watkins, 1979), chi-square (McCall, 1975), kappa statistics (Cohen, 1960), and item-objective congruence (Thorn & Deitz, 1989) - have appeared useful.

Protocol Analysis

Messick (1989) suggests a few methods to analyze the processes underlying test items or task performance. Protocol analysis is regarded as the most prominent among these methods. It is noteworthy that protocol analysis is seldom used by occupational therapists in validation studies. Analyses of testing processes contribute to the substantive and structural validity of the test interpretation and use. In protocol analysis, individuals are asked to think aloud during task performance or to describe retrospectively the procedural steps they employed. The verbal reports are then analyzed by some form of discourse analysis. Protocol analysis can be applied to performance assessments such as assessing the ability of clients to perform self-care activities, and self report assessments such as assessing degree of the satisfaction and role performance of clients.

Protocol analysis was developed from an information-processing theory in cognitive psychology (Anderson & Bower, 1973; Simon, 1979). The theory postulates a transformation process which underlies protocol analysis (Erisson & Simon, 1980). First, the theory suggests that a cognitive process can be depicted as a sequence of internal states which are transformed successively from a series of information processes. The theory goes on to postulate that each successive state can be largely described in terms of a small number of information structures, or chunks. These pieces of information are attended to or are available in the short term memory store (STM) of an individual, and can be retrieved through verbalizations. Based on this theoretical framework, the verbal protocols of individuals can reflect the ongoing cognitive processes when they are working on different tasks.

The ideal method to gain information about respondents' information processes, according to Ericsson and Simon (1985), is to ask the respondents to "think aloud" while they are working on tasks. Instructions, such as "I want you to tell me what you are doing or what you are thinking about while attempting to prepare a menu for tonight's dinner" or "...tell me what you are thinking about while you are deciding the importance of you preparing a meal for yourself", are given at the beginning of or during the assessment. An alternative method is to ask the respondents to report their thought sequences verbally just after tasks have been completed. Instructions such as "how did you complete this menu for tonight's dinner?" or "how did you decide that preparing a meal is very important to you?" are given immediately after the clients finish the tasks. According to Ericsson and Simon (1980), the retrospective strategy is less desirable because of the possibility of clients reconstructing their thoughts. The longer the time lag the less reliable the report because the clients' short term memory is not being accessed directly.

In both of these verbal reporting methods, clients are asked to verbalize their thoughts as they emerge without trying to explain, analyze, or interpret those thoughts. Interviewers can use both directed and undirected probing to elicit more information from the clients. The choice of probes should be relevant to the context of the tasks and the type of evidence that researchers want to establish. Interviewers can record the verbal report by writing verbatim or audiotaping. Qualitative content analysis after systematic coding of the verbal reports is a common strategy used to analyze the data gathered. Evidence on the consistency between the processes underlying the test items and test behavior produced by clients and those postulated by the test construct on which the clinical instrument is built can then be examined. There is one general drawback with this method: protocol analysis has the potential to produce inconsistent information (Ericsson & Simon, 1980). First, the information revealed in the verbal report may not be identical with the information in STM because of retrieval errors made by the respondent. Second, the respondent may have a tendency to fill out and generalize incomplete memories, particularly in recall situations. Third, this strategy is not suitable for tasks requiring rapid mental processes. As a consequence, clinicians utilizing protocol analysis should make use of their clinical judgement, observational skills, and ability to probe or question to ensure the quality of information gathered.

Multi-trait Multi-method Matrix

The multi-trait multi-method matrix is another well-known method used to gather evidence on the substantive and external components of test validation (Messick, 1989). The method was developed by Campbell and Fiske (1959, 1967). It is built on the concept of convergent and discriminant validation. This method is said to be superior to a monotrait-monomethod approach because:

> Any single operation, as representative of concepts, is equivocal.... The addition of a second viewpoint, as through binocular parallax, greatly

> reduces this equivocality, greatly limits the constructs that could jointly account for both sets of data. (Campbell & Fiske, 1967, p.129)

The multitrait-multimethod technique (Campbell & Fiske, 1959) is used to determine the extent to which the scores of an instrument are correlated with the scores from other tests of varied test constructs and test formats. Various hypotheses are formulated: 1) the correlations of the scores on the validated instrument with the scores from tests measuring similar traits but by different formats/methods are high (known as convergent validity); 2) the correlations of the scores on the validated test and the scores from tests of similar traits using different formats/methods are higher than the correlations between those measuring different traits using the same formats/methods (known as discriminant validity) and are higher than those measuring different traits by using different formats/methods; 3) the pattern of correlations among traits should be similar for the same and different methods; 4) the correlations between the scores of a test and itself (known as reliability coefficients), including the instrument of validation, should yield the highest values. Campbell and Fiske suggest that such a method is good for a smaller sample size. In all cases, it is necessary to include at least two independent traits, each measured by at least two methods.

In the occupational therapy literature, the use of the multitrait-multimethod technique has been rare. In a partial approach, Rothman, Hedrick, and Inui (1989) used the concepts of convergent and discriminant validity to study the relationship of the Sickness Impact Profile with the health indicators such as the Index of Activities of

Daily Living, Barthel Index, Life Satisfaction Index Z, and the Philadelphia Geriatric Centre Morale Scale. Their findings suggested high correlations between the Sickness Impact Profile, and the Barthel Index (*r*= .95) and ADL index (*r*= .74). In contrast, the correlations between the Sickness Impact Profile and two other psychological well-being measures were low and in an opposite direction (*r*=-.31 with the Life Satisfaction Index Z and *r*=-.40 with the Philadelphia Geriatric Centre Morale Scale). In the analysis of the comprehensiveness of the Sickness Impact Profile, Rothman et al. concluded that some of the items in the Sickness Impact Profile present the dimensions of behavior assessed by the two functional assessment instruments, which is convergent validity. However, they did not provide detailed comments on the discriminant validity, that is, the difference between the constructs of "sickness" and "satisfaction and morale". Further explanations on the discriminant validity would help illuminate the uniqueness of the instrument and the unshared variance due to differences in constructs.

Conclusion

Messick's (1989) concept of validity and validation provides a unified, comprehensive, and in-depth model to guide evaluation of a test score and its interpretation. However, Messick's model of validation has not been applied to the practice of occupational therapy and other rehabilitation professions. Instead of making significant changes to the notion of validity, Messick's model redresses and integrates various concepts of several prominent theorists in the measurement field. Its approach to validation is a summation of various strategies and statistical techniques that have been frequently utilized in previous validation studies. The strengths of Messick's model of validation are that 1) it unifies the concept of validity and conceptualizes the evidence of validity of test results are gathered from various sources; 2) it brings the consequences of test interpretation and use into the notion of validity which has previously remained untouched; and 3) it balances the focus of validation on both internal and external properties of an assessment tool which has rectified the ill-formed concept of having "criteria" as the gold standard.

Validation of the COPM in this study used Messick's model as a blueprint to guide the process and strategies in gathering evidence from different sources. Equal emphasis was placed on each of the various components including content-related, substantive, structural, and, criterion related evidence of validity. The major weakness of this study is the lack of consequential evidence. Due to the limited scope inherent in any single study, the validation process only concentrated on two cells of Messick's four-cell model.

CHAPTER III

METHOD OF INVESTIGATION

Introduction

This chapter describes the strategies for collecting evidence of construct validity of the COPM. It begins with a description of the sample and the recruitment method. This is followed by a description of the instrumentation, the design of the validation process and the data collection procedures.

Sample and Sampling Method

The population with which evidence of construct validity of the COPM was established using adult clients with physical disabilities who were receiving occupational therapy programs through the Glenrose Rehabilitation Hospital in Edmonton, Alberta. The sample consisted of clients selected from two diagnostic groups: restorative orthopedics (Orthopedic) and cerebral vascular accidents (Stroke). The decision to use diagnosis as a grouping variable was based on the assumption that clients in the same diagnostic group were mostly likely to manifest similar forms of physical disabilities (World Health Organization, 1980) and, hence, similarity in their hospital length of stay, physical capabilities, clinical manifestations, treatment rationales, and occupational therapy programs could be expected (Hopkins & Smith, 1988; Trombly, 1988). Nevertheless, clients from the same diagnostic groups are likely to possess different personal characteristics such as social, cultural, and financial backgrounds. Classification of clients based on diagnosis enhanced the within-group homogeneity. It was hoped that this homogeneity would yield data that were unconfounded by a wide range of extraneous variables associated with the disability of these clients, thus allowing stronger inferences to be made from the results of the study (internal validity). The use of only two diagnostic groups in the present study also allowed the process of data collection to be standardized within groups, yet allowed meaningful between-group comparisons between the orthopedic and stroke clients.

Another purpose of selecting the orthopedic and stroke diagnostic groups was to increase the generalizability of the evidence generated in this study. The more the differences (as mentioned previously) between clients in the two groups, the greater the between-group heterogeneity. Hence the higher the possibility of generalizing the results of the study to clients with other kinds of disabilities. Given the suggestion that the COPM possessed universal utility for all disability groups (Law et al., 1991), using two diagnostic groups with very different clinical and functional characteristics was aimed at challenging this claim.

The following selection criteria were used to identify potential clients for the study. The clients:

1) were admitted to the hospital and belonged to either an orthopedic or stroke group;
2) were over 17 years of age;

3) were receiving occupational therapy at the period of data collection;

4) had no cognitive or communication disorders that affected their judgement and problem solving abilities (as screened by the Neurobehavioral Cognitive Status Examination); and

5) voluntarily gave their consent to participate in the study.

Clients who had either significant cognitive or communication disorders as identified by the screening test (stated in criterion 4) or did not complete the programs provided by the hospital's occupational therapy department were excluded from this study.

The sample size for each diagnostic group was set at 30 yielding a total sample size of 60. Although neither power and effect size analyses had been done, 30 clients per group is generally accepted as an adequate sample size in clinical research (Portney & Watkins, 1993). Selection of clients followed a non-probability convenience sampling method (Smith & Glass, 1987). This sampling method was used because the number of clients was comparatively small in each of the hospital's diagnostic groups. The lack of full freedom in selecting and scheduling clients for conducting the assessments and interviews for the study due to clients' commitments to their own clinical routines also limited the choice of other sampling methods such as systematic and stratified random sampling.

All clients with the diagnosis of either orthopedic problems or stroke and who were referred for occupational therapy during the period of this study were screened for their suitability. The screening of clients using the Neurobehavioral Cognitive Status Examination (Northern California Neurobehavioral Group, 1988) and other inclusion criteria was carried out for each client by his/her case therapist. Once the clients were identified as suitable for the study, the case therapists scheduled assessment sessions and implemented the various assessments and interviews in accordance with the study protocol.

Instrumentation

In this section, the instruments and their psychometric properties used for establishing the criterion-related validity in the present study are reviewed. The characteristics of the Canadian Occupational Performance Measure were discussed in Chapter II and will not be repeated here. The instruments are described in the following order: the Neurobehavioral Cognitive Status Examination, the Role Checklist, the Satisfaction with Performance Scaled Questionnaire, the Klein-Bell ADL Scale, and the Functional Independence Measure.

Neurobehavioral Cognitive Status Examination

The Neurobehavioral Cognitive Status Examination (Northern California Neurobehavioral Group, 1988) is a screening tool. It was used to identify clients who were cognitively impaired. According to the Test Manual for the NCSE, the criterion cut-off scores are set for each of ten sub-scales. Clients who scored below the cut-off scores, indicating obvious cognitive impairment, were excluded from the study. This test generally takes about 15 to 30 minutes to administer (Cammermeyer & Evans, 1988).

The NCSE is designed to assess the intellectual functioning of clients in five major ability areas: Language, Constructions, Memory, Calculations, and Reasoning (Northern California Neurobehavioral Group, 1988). Attention, Level of Consciousness, and Orientation are assessed independently. Language has four separate subsections: Spontaneous Speech, Comprehension, Repetition, and Naming. Reasoning has two subsections: Similarities and Judgement.

The instrument uses a screen and metric approach. The results of the test are presented in terms of cognitive status profiles with reference to the various ability areas assessed. In four of the five ability areas (the Memory subtest is excluded), a client is first presented with a "screen" item which taps a criterion-referenced level of certain ability. If the client passes the screen item, the ability involved is assumed normal and no further testing is done in that subtest. If the client fails the screen item, the examiner then administers the "metric" items of that subtest which were constructed with increasing difficulty. Performance on the metric items of a subtest determines whether and to what degree a particular ability is impaired.

The NCSE has been normed for three different age groups (Kiernan, Mueller, Langston, & Van Dyke, 1987). According to the standardization data, specific cut-off scores are assigned to different subtests. Scores above the cut-off (e.g., a score of 10 on the Memory subtest) are classified as within the average range. Scores below the cut-off scores are further classified into mild, moderate, and severe cognitive dysfunction.

Reliability of NCSE

The review of the literature on the NCSE did not provide much information on the reliability of the instrument. Kiernan et al. (1987) suggested that the usual reliability criteria did not apply to the NCSE based on the fact that healthy clients should perform almost perfectly on all the subtests of the instrument. They further commented that the test and retest method was not meaningful due to this ceiling effect. The authors also specified the inappropriateness of using a split half method because of the small number of items that the NCSE contains. This was regarded as a drawback when using the instrument in this study. However, in view of its high sensitivity to detect clients with cognitive impairment, the screening tool was used in this study.

Validity of NCSE

The NCSE was found to be more sensitive in the detection of cognitive dysfunction and had a significantly lower false-negative rate than the Mini-mental State Examination (MMSE) and the Cognitive Status Examination (CCSE) (Schwamm, Van Dyke, Kiernan, Merrin, & Mueller, 1987). The false-negative rate of NCSE was reported as 7% in identifying clients with central nervous system lesions that were confirmed by computer tomography, magnetic resonance imaging, or biopsy. The sensitivity of the instrument was derived from two features of its design: the use of independent tests to assess skills within five major areas of cognitive functioning and the use of graded tasks within each of these cognitive domains.

Use in this Study

The NCSE was used in the present study as a screening tool to identify clients who were cognitively impaired. The assessment was conducted by clients' case therapists. Clients were not selected for the study if they scored below the cut-off scores as defined by the test.

Role Checklist

The Role Checklist is used to measure the perceived role incumbency of the clients and the degree to which each role is valued by clients (Oakley, Kielhofner, Barris, & Reichler, 1986). The Role Checklist was built on the model of human occupation (Kielhofner, 1985) which shares a similar construct with the model of occupational performance.

The Role Checklist uses a paper and pencil format (Appendix II). It elicits the client's perceived role incumbency in ten roles along a temporal continuum (Oakley et al., 1986). The instrument requires about 15 minutes to complete and may be done with or without one-on-one supervision from therapists.

Reliability of Role Checklist

To estimate the test and retest reliability of the Role Checklist, Oakley et al. (1986) administered the instrument to 124 subjects on two separate occasions (either 1-4 weeks or 5-8 weeks). The percentage of agreement of subjects' responses between the two occasions for each of the ten roles ranged from 76% to 94% (n=87) for the 1-4 week group and 79% to 93% (n=37) for the 5-8 week group.

Validity of Role Checklist

Oakley et al. (1986) presented little evidence with respect to the validity of the Role Checklist. The authors used a panel composed of occupational therapy graduate students, faculty members, and practising therapists to evaluate the relevance of the proposed role taxonomy. No detailed results of the review were reported. The instrument was also field tested on a group of clients in an inpatient psychiatric program. Again, results of the field study were not presented. Oakley et al. noted that the comments gathered from the panel review and field testing had been considered and subsequent changes had been made to select appropriate roles and modify the role definitions in the instrument.

Use in this Study

The COPM (Problem Definition and Problem Weighting) and the Role Checklist use different methods to measure life roles and role expectations of the clients; i.e., a semi-structured interview versus paper-and-pencil tasks. Convergent validity was expected to be established by comparing the results of the two instruments. The evidence gathered through content analysis of the data would therefore contribute to the substantive validity of the COPM.

Satisfaction with Performance Scaled Questionnaire

The Satisfaction with Performance Scaled Questionnaire (Yerxa, Burnett-Beaulieu, Stocking, & Azen, 1988) is a 46-item questionnaire used to measure clients' satisfaction with their performance of independent living skills (Appendix III). According to Yerxa and colleagues, independent living skills are operationalized in two dimensions - tasks related to home management and tasks related to social/community problem solving. The SPSQ consists of a 24 item home management subscale and a 22 item social/community problem solving subscale. The home management subscale has 24 common household tasks, such as "make a bed" and "put clothes on hangers". Clients are required to quantify the proportion of time (0, 25, 50, 75 and 100%) that they feel satisfied with the way they have performed the listed tasks. The same rating method is used for the social/community problem solving subscale with items such as "socialize with other persons" and "find and use other social activities".

Reliability of SPSQ

Yerxa et al. (1988) reported that the internal consistency of items, based on the responses from 50 subjects, was .97 for the home management subscale and .93 for the social/community problem solving subscale (a split-half method of estimation).

Validity of SPSQ

Content validity of the SPSQ was established by correlating the scores on the SPSQ with the scores on overall life satisfaction. Clients' overall life satisfaction was measured by a single item by which clients rated their satisfaction on a 3-point scale. The product-moment correlation coefficient was .62 ($p < .01$) for the social/community problem solving subscale and .44 ($p < .05$) for the home management subscale (Yerxa & Baum, 1986). Significant relationships were also found between satisfaction with an activity and self-perceived independence in performing the activity. Evidence on the construct validity of the SPSQ was collected in a study comparing 15 community based clients with spinal cord injuries with 12 non-disabled cohorts matched for age, sex, ethnicity, and community of residence. The clients with spinal cord injuries scored significantly lower than the non-disabled cohorts on both scales of the SPSQ.

Use in this Study

The SPSQ was not built on a client-centred model and does not take a respondent's life roles and role functions into account. Moreover, it does not include an evaluation of a client's performance. Hence, convergent relationships were expected between clients' scores on the SPSQ and their scores on the Satisfaction sub-scale of the COPM. In contrast, clients' scores on the SPSQ were predicted to be relatively unrelated to the scores on the Performance subscale of the COPM (discriminant validity).

Klein-Bell ADL Scale

The Klein-Bell ADL Scale (Klein & Bell, 1979) measures clients' performance in terms of the degree of independence with reference to a well defined set of self-care activities (Appendix IV). There are a total of 170 tasks classified under six areas of self-care activities. The six subscales are: eating, dressing, mobility, bathing, elimination, and communication. The 170 tasks were selected based on expert panel reviews composed of health care professionals. This instrument was not built on a client-centred model, and it uses performance and observational formats. According to the Test Manual (Klein & Bell, 1979), clients to be assessed are asked to perform pre-determined self-care tasks while the therapists observe and rate their performance. Responses to items use a dichotomous scale of "able" and "unable" to perform.

Reliability of the Klein-Bell ADL Scale

The inter-rater reliability was estimated using six pairs of raters. Across all items on 20 patients, 92% agreement was achieved (Klein & Bell, 1979).

Validity of the Klein-Bell ADL Scale

The only evidence on predictive validity was provided by the Klein and Bell. A Pearson product moment correlation coefficient of -.86 ($p < .01$) was reported between scores on the instrument and the number of hours per week a person required assistance during a five- to ten-month period after discharge (Klein & Bell, 1982).

Use in this Study

For the purpose of this study, the Klein-Bell ADL Scale was modified to form a performance checklist. Case therapists completed the checklist based on their clinical observations of and professional judgement about clients' functional performance. The checklist took about 30 minutes to complete.

The Klein-Bell ADL Scale had dual relations with the COPM. The Klein-Bell ADL Scale has a similar construct (measuring clients' performance in self-care activities) but different methods (expert judgement versus self-report) when compared to the COPM. Convergent validity was expected between the scores of the clients on the Klein-Bell ADL Scale and the Performance subscale scores on the COPM. In contrast, divergent validity was expected by the scores on the Klein-Bell ADL Scale and the scores on the Satisfaction subscale of the COPM.

The Functional Independence Measure

The Functional Independence Measure (FIM) (Granger, Hamilton, & Sherwin, 1988) is a well known instrument used in many clinical settings to measure clinical outcomes and effectiveness of rehabilitative services (Appendix V). According to Granger and his associates, the FIM was constructed to quantify severity of disability, functional status, and burden of care of clients receiving rehabilitation services. The instrument is based on 18 critical tasks and attributes that clients need to accomplish in order to function independently. It adopts a multidimensional concept of functional capability (Granger, Hamilton, & Sherwin, 1988; Shah & Cooper, 1993). The 18 items are organized under six categories of general functions of clients, namely self-care, sphincter control, mobility, locomotion, communication, and social cognition (Hamilton, Granger, Sherwin, Zielezny, & Tashman, 1987). All the items are divided into two subscales, namely the motor subscale (the 13 items of the first four categories) and the cognitive subscale (the 5 items of the communication and social cognition categories). Rehabilitation professionals observe and rate their clients on each item with a seven-point scale ranging from Complete Independence (7 points) to Total Assistance (1 point) (Center for Functional Assessment Research, 1991). Motor and cognitive subscale scores are yielded by summing the scores of items in the two subscales.

Reliability of FIM

The test-retest reliability and internal consistency are not reported in the Test Manual. The authors of the instrument commented that these properties were less important because of the intended use of the instrument (essentially for program evaluation), the nature of clients (known to exhibit slow change on key variables), and the number of items (relatively few items are included in the instrument) (Hamilton et al., 1987). In two reliability studies conducted in 1993, results showed that the inter-rater reliability of administering the FIM was satisfactory. Heinmann et al. (1993) reported that the inter-rater agreement of FIM total scores, in terms of intra-class correlation coefficient, was 0.90 or greater; the estimates for the six FIM category scores were all above 0.75; and at least 14 of the 18 FIM items equalled or exceeded a Kappa score of 0.45. Fricke, Unsworth and Worrell (1993) in their study of inter-rater reliability reported values for intra-class correlation coefficients that ranged from 0.75 to 0.94.

Validity of the FIM

In the original article written by the authors of the instrument (Hamilton et al., 1987), a qualitative approach was adopted to establish evidence on the construct validity of the FIM. The authors used the term "face validity" to refer to the relevance and representativeness of the items to the construct of functional independence. The results from an expert panel review were reported as the "face validity" appeared to be "good". In a more recent article, the authors applied

Rasch analysis to establish more evidence on the construct validity of the FIM (Heinemann et al., 1993). Rasch analysis revealed that unidimensionality of the items exists when the 18 items are grouped into the two subsets of motor (13 items) and cognition (5 items). When the data were divided according to each impairment group, the standard deviations of item difficulties, average mean squares of items (3% of 1.0), and typical standard error of items (0.01 Logits) indicated acceptable item-model fit and the consistency of the items within the motor and cognitive subscales. Factor analysis also revealed a satisfactory two-factor model among the 18 items (Heinemann et al., 1993). The 13 motor items were found to account for 95% of the variance of the motor factor while the 5 cognitive items accounted for 92% of the variance of the cognitive factor.

Use in this Study

The FIM was not built on the client-centred and occupational performance models. The content of FIM and the COPM are very different. Analysis of the test content suggested that convergent validity was expected between clients' motor subscale scores on the FIM and the Performance subscale scores on the COPM. In contrast, clients' cognitive subscale scores on the FIM were expected to be relatively unrelated to either the Performance and Satisfaction subscale scores on the COPM.

In the present study, the scores on the FIM of clients in the stroke group were extracted from the clients' medical report in the hospital. The FIM was administered to clients in the orthopedic group by the participating therapists both in the initial and pre-discharge assessments.

Design of the Validation Process

This study used Messick's (1989) model of validation as its framework to guide the identification and collection of data. The evidence gathered was limited to the evidential basis of interpretation and use of the COPM, as described in the previous chapter. Based on Messick's framework, the objectives of the study were to investigate:

1) the content-related evidence in terms of the relevance and representativeness of the COPM's test domain and the construct domain defined by the occupational performance model;
2) the substantive component of construct validity by comparing the test behaviors actually exhibited by the clients during administration of the COPM and the behaviors thought to be measured by the COPM;
3) the structural component of construct validity by examining the relations among items, sub-scale scores, and total scores of clients on the COPM, and comparing the results of the assessment to the corresponding results speculated by the occupational performance model; and
4) the criterion-related evidence by examining the relationships between the sub-scale scores of clients on the COPM and their scores on the other tests measuring similar and different traits, and by similar and different methods.

In order to achieve the objectives as stated, data were gathered by administering the COPM and four other instruments to clients in the two groups on two assessment occasions, namely the Role Checklist, Klein-Bell ADL Scale, Satisfaction with Performance Scaled Questionnaire, and Functional Independence Measure. The first occasion occured at the initial assessment on admission to the hospital and the second just prior to discharge. The underlying rationale behind the administrations of the COPM and other instruments at these two occasions was that the "COPM scores are meaningful for comparative purposes, that is change from assessment to re-assessment" (Law et al., 1991, p.17). As a consequence, the two sets of scores obtained on these assessment occasions were compared to identify any differences and plausible trends in the assessment scores. In this study, major efforts were made to study the intra-individual relations among different variables on the initial assessment, pre-discharge assessment, and between the two assessment occasions.

Stability of the COPM

Besides the four objectives listed in the previous section, there was a fifth objective. That objective was to investigate the stability of clients' sub-scale and total scores on the COPM between different times. The test-retest method was employed in an attempt to establish the stability of the clients' scores on the COPM between an initial assessment and a retest assessment occasion. It was thought that the method would give an indication of the reliability of the instrument.

The test-retest method was chosen over the method of internal consistency because the COPM has comparatively few items (only 5 activity items) in the ratings of Importance, Performance, and Satisfaction. Moreover, the activity items (the activities client identified as difficult to perform) were not intended to be comparable across clients. For example, one client may have difficulties in gardening while another client may find communicating with others difficult. The differences in activities identified by clients were attributed to the fact that problems with functional tasks of each client are multi-factorial in origin (Christiansen, 1991a; Reed, 1984). Since items differed across people, computation of internal consistency was not meaningful.

In the present study, the test-retest method was supposed to provide estimates of reliability in terms of stability of clients' scores obtained on two assessment occasions at different times. It was planned that 30 clients, with 15 from each of the diagnostic groups, would be selected for retest from the original sample. After the initial assessment, the clients selected for retest were assessed again by the same therapist using the COPM (including Problem Definition, and ratings of Importance, self-perceived Performance and Satisfaction with performance).

The time lapse between the test and retest was a clinical judgement with the aim of reaching a compromise between two factors - minimizing the effects of the initial testing upon clients and at the same time minimizing the changes in clients' functional status. The administration of the COPM to the clients in the initial assessment would produce the unavoidable memory and practice effects upon the clients. These effects would inflate the stability of clients' responses since clients could reproduce their responses based on their memory and experience of the ways they responded in the initial assessment. One way to minimize the memory and practice effects is to increase the time lag between the initial assessment and retest of the clients. However, an increase in time lag for administering the retest would deflate the stability of clients' responses due to the changes within the clients. These changes could be attributed to the treatment effects of rehabilitation programs upon the physical disabilities of the clients and/or client's natural recovery. An ideal way to avoid the influence of within client changes between initial assessment and retest is to repeat the COPM immediately after the initial administration. However, the problems with this strategy are obviously the undesirable memory and practice effects upon the clients since the COPM does not have a parallel form version.

In view of these two opposing factors, the clinical utility of the COPM becomes an important element in deciding the time lag between initial assessment and the retest of clients. Clients receiving rehabilitation programs in hospital usually show rapid improvement in their functional status. An outcome measure such as the COPM should theoretically be administered to clients as frequently as possible in order to capture those changes. The COPM is a clinical instrument and its clinical utility (rapid repeated measures) supported the choice of a shorter duration between the initial assessment and retest of the clients in the test-retest stability study. Since clients in the orthopedic group have a much shorter average length of stay than those in the stroke group (ten days versus six weeks according to the hospital statistics), the difference should be reflected in the time line between initial assessment and the retest. The clients in the orthopedic group were retested within three days after the initial administration of the COPM, while the clients in the stroke group were retested within seven days after the initial assessment. This schedule was a compromise between various undesirable factors which are believed to lower the interpretability of the test and retest stability estimate of the COPM. Consequently, as the results will show, the reliability estimates calculated had no creditability. Nonetheless, evidence gathered in this part of the study provided useful indications of the stability of clients' responses on the COPM over a short period of time. The results also contributed to a better understanding of the meaningfulness of the instrument.

Data Collection Procedures

The data collection process was divided into three stages. Different stages were designed to gather evidence on the different components constituting construct validity of the COPM. The three stages were: development of testing protocols, the panel review (content-related evidence), and the clinical assessment processes (substantive, structural and criterion-related evidence, and test-retest reliability).

Stage 1 - Development of Testing Protocols

Four occupational therapists working in the orthopedic and stroke teams at the Glenrose Rehabilitation Hospital were recruited to assist with the development of the testing protocols. These protocols included the formulation of assessment procedures and schedules with respect to the initial assessment (pre-test), retest, and the pre-discharge assessment (post-test). The purpose of involving the therapists at the initial stage was to ensure minimum disruption to the clinical procedures when the assessment protocols of this research were merged into those regularly used in the hospital. Standardization of the assessment protocols was also important to gain better control of confounding variables that might reduce the quality of the evidence gathered. Examples of some of these undesirable effects were variations in the assessment procedures, the differences in time of the two assessment occasions, and differences between raters administering the assessments. The assessment protocols developed for this study were designed to minimize those factors mentioned and meet the standards of practice set out in the departmental procedure manuals of the Occupational Therapy Department.

Significant difficulties were met in trying to unify the time line for administering the assessments to the clients in the two diagnostic groups in terms of particular days and weeks because of variation in clients' clinical condition, rehabilitation programs received, and discharge criteria for the two groups. As a result, clients' schedules of assessment were defined in terms of different treatment stages.

The rationale for setting two different assessment schedules for the clients in the two diagnostic groups was based on the fact that there was a significant difference in the average length of stay of clients in the two groups. According to the statistics provided by the hospital, the average length of stay of clients in the orthopedic group is ten days, whereas for the stroke group the average length is six weeks. The two assessment schedules were set in proportion to the respective lengths of stay of the two groups with the idea that the COPM would reflect clients' behavior at particular stages of rehabilitation, and hence allow inter-group comparisons.

For clients in the orthopedic group, the Initial Assessment was administered within three days after requisition for occupational therapy consultation was received by the case therapists. For the stroke group, the corresponding time was 14 days after receiving client referrals. In case of Re-assessment, clients selected for the retest were reassessed within three days after the date of their initial assessment if they were in the orthopedic group and seven days if they were in the stroke group. Lastly, Pre-discharge Assessment was conducted within three days for the orthopedic group and seven days for the stroke group before the exact or tentative date of discharge.

Stage 2 - Panel Review by Clinical Experts

Using the concept of expert review, a nine-member panel was set up to establish evidence on the content-related and structural validity of the COPM. Of the nine panel members, seven were clinical specialists working in different treatment programs at the Glenrose Rehabilitation Hospital. Each of them had experience in administering the COPM and working with orthopedic and/or stroke clients. The remaining two members were a clinical measurement expert and an occupational therapy faculty member from the Department of Occupational Therapy, University of Alberta. They were selected because of their expertise in clinical measurement and occupational therapy.

All of the seven clinical specialists received a three-hour training session prior to the panel review session. The training included introduction to the occupational performance model, test construction theories, psychometric characteristics of assessment instruments, and functions of the panel review. For the other two members, informal seminar sessions were organized individually to provide information on the aspects that they were less familiar with (e.g., the theoretical model of occupational performance and the administration procedure of the COPM for the measurement expert and Messick's model of validation for the occupational therapy faculty member).

A questionnaire was designed to guide the review process of the COPM during the panel review (Appendix VI). Panel members were required to complete the questionnaire. The questionnaire was divided into three parts: 11 items were related to the content of the COPM, 8 items were related to structural aspects of the COPM, and 19 items addressed utility and procedural issues. The content review section was used to evaluate the degree of relevance and representativeness of the items of the COPM in relation to the construct of occupational performance as defined in Chapter II. The structural review section was focused on the degree to which the rating scales and the scoring system of the COPM represented the structural relations of the different variables postulated in the construct of occupational performance. The utility and procedural review section was used to assess the practicality and universal application of the COPM as claimed in the Test Manual, and to determine the strengths and weaknesses of the administrative procedures of the COPM.

The content of the questionnaire was based on the evaluation models of clinical instruments suggested by Christiansen and Baum (1991), Law (1987), and

Loevinger (1957). Items were constructed for each area of review to collect opinions (both close-ended and open-ended) from the panel members. All of the close-ended items in the questionnaire were rated using a five-point ordinal scale: Excellent (5), Very Good (4), Good (3), Fair (2), and Poor (1). Panel members were to evaluate the COPM by rating each item with reference to the ordinal scale. For some of the items, open-ended questions were also presented. Panel members were asked to rate the item and then comment on particular aspects of their evaluation.

The panel members received the questionnaires separately and were given one hour to complete the review. During the review, panel members were reminded not to communicate with one another so as to ensure maximum independence of the opinions gathered. The questionnaires were collected after the panel members had completed their evaluation.

Stage 3 - Data Collection Process

Sixty clients were to be assessed with the COPM and other instruments on two different occasions at the initial and pre-discharge assessments. Eight qualified occupational therapists were recruited to collect data from the 30 clients in each of the orthopedic and stroke groups.

As noted earlier, in order to gather different sources of evidence on the construct validity of the COPM, in addition to the COPM, five other instruments were administered to all the clients at the initial assessment phase (Table 3.1).

Table 3.1

Assessment Protocols of the Clients in the Three Assessment Occasions

Initial Assessment (Total Sample) N = 60	Re-assessment (For Stability Study) *n* = 30	Pre-discharge Assessment (Total Sample) N = 60
NCSE (1)	----	----
COPM	COPM	COPM
Role Checklist	----	----
Klein-Bell ADL Scale	----	Klein-Bell ADL Scale
SPSQ (2)	----	SPSQ
FIM (3)		FIM

Notes: (1) NCSE: Neurobehavioral Cognitive Screening Examination

(2) SPSQ: Satisfaction with Performance Scaled Questionnaire

(3) FIM: Functional Independence Measure

These were the Neurobehavioral Cognitive Screening Examination (NCSE), Role Checklist, Klein-Bell ADL Scale, Satisfaction with Performance Scaled Questionnaire (SPSQ), and Functional Independence Measure (FIM). Descriptions and psychometric properties of each instrument were presented in the instrumentation section of this chapter. In the re-assessment and pre-discharge assessment, intra- and inter-client comparisons were achieved by repeated administrations of selected instruments (see Table 3.1).

All three assessment occasions were incorporated into the orthopedic and stroke occupational therapy programs. For the stroke group, the FIM was administered as part of the routine assessment procedure used by the stroke team (prior to this research). Therefore, the FIM data were obtained by the case therapists directly from the medical records of the clients in the stroke group. The NCSE was administered to the clients for screening purposes. This was then followed by the COPM. The sequence of administering the Role Checklist, SPSQ, Klein-Bell ADL Scale, and FIM to the clients was largely dependent on the convenience of participating therapists.

To facilitate these assessments, the participating therapists attended a series of training sessions on the standard procedures of administering the COPM, Role Checklist, SPSQ, Klein-Bell ADL Scale, and FIM. The training sessions included a two-hour seminar presented by the researcher on the administration of the COPM to ensure standardization of the testing procedures among the raters (Appendix VII). Detailed information was provided on the assessment schedules, the consent form (Appendix VIII), the client information sheet (Appendix IX), subject's data base (Appendix X), and other relevant background information on the model of occupational performance.

Clients were approached by their case therapists and invited to participate in the study. The client consent form and client information sheet were shown and explained to each client. Clients who chose to participate signed the consent form. The total assessment time for each client was estimated to be less than three hours. The initial and pre-discharge assessments took about two hours to complete, and the re-assessment (explained in detail in the following paragraph) took about half an hour to complete.

From the total sample of 60, 30 clients were to be selected for the re-assessment. Re-assessment occured within three days from the date of initial assessment for the orthopedic group and seven days for the stroke group. The method of selecting the sub-sample was by convenience sampling on a first-come-first-served basis. The COPM was administered to each of the selected clients for the second time by the same therapist as in the initial assessment occasion. Besides the COPM, the clients were interviewed by the therapists. A set of leading questions was constructed to prompt the clients to describe retrospectively the processes they engaged when they responded to the questions and ratings when assessed with the COPM. The interviews with the clients were semi-structured. The function of the leading questions was to set the parameters for the therapists to facilitate their exploration of the mental processes that the clients used during the test. Some of the leading questions were:

> How did you know that you would have problems with "getting in and out the bath-tub when you go home"?;
>
> How did you come up with the rating of 8 to "the importance of shopping to you"?; or
>
> What was in your mind when you rated your performance of "putting on a blouse" as a 7?

Data gathered from the interview were then used as evidence related to the substantive validity of the COPM.

CHAPTER IV

RESULTS AND BACKGROUND ANALYSIS

Introduction

Chapter IV provides an analysis on the characteristics of the clients who participated in this study. This is followed by the results of data collection with special focus on the COPM and other clinical instruments.

Clients and Assessment Results

The original sample size of the study was targeted at 60, with 30 clients in each of the orthopedic and stroke groups. However, during the period of data collection, the hospital from which the sample was selected experienced many major changes to its clinical and management structures that hampered the progress of the study. In addition, the functional status of potential clients in the stroke group was found to be much lower than anticipated. Participating clinicians encountered difficulties in selecting clients who met the inclusion criteria (especially passing the NCSE screening test). As a result, only nine clients in the stroke group were selected to complete the assessment protocols. Together with the 30 clients assessed in the orthopedic group, the total number of clients was 39.

The unforseeable changes in the hospital and in the stroke client groups led to a substantial reduction of the sample size. This in turn resulted in changes to some of the strategies used in analyzing and presenting evidence on the construct validity of the COPM. These changes will be discussed together with the results in later chapters.

Altogether, nine occupational therapists including the researcher participated in collecting the data. The entire data collection process took about eight months to complete.

The Sample

Among the total sample of 39 clients, 11 (28.2%) were male and 28 (71.8%) were female. The proportions of male and female clients were different in the orthopedic and stroke groups. Ten (33.3%) out of 30 clients in the orthopedic group were male, whereas only one (11.1%) out of 9 clients in the stroke group was male. The mean age of the sample was 64.5 years. The mean ages of clients in the orthopedic and stroke groups are presented in Table 4.1.

The majority of the clients were Caucasian, with only one Asian and one Native Indian in the stroke group. The average length of stay for all clients was 23.3 (*SD*=13.4) days, although there were differences between the clients in the orthopedic and stroke groups. For the orthopedic group, the average length of stay was 19.2 (*SD*=10.0) days. In contrast, the average length of stay in the stroke group was 36.9 (*SD*=14.7) days.

Table 4.1

Mean Age of Clients in the Orthopedic and Stroke Groups

		Age of Clients (in years)	
Groups	*n*	M	SD
All	39	64.5	15.1
Orthopedic	30	65.3	14.7
Stroke	9	62.0	17.1

Results of COPM

The COPM was administered to all of the 39 clients. The average number of days between the initial and pre-discharge assessments was 19.8 days (SD=22.1 days). The interval between assessments in the two occasions was inflated by one client in the stroke group who was discharged from the hospital and continued receiving out-patient occupational therapy. For that particular client, the interval between the two assessments was 141 days. Removal of this client from the group greatly reduced the mean number of days between assessments to 16.6 days (SD=9.7 days). For the orthopedic group, the mean was 15.4 days (SD=9.9 days) with a maximum of 46 days and minimum of five days. As anticipated, the stroke group had a longer interval between assessments with the mean 20.9 days (SD=7.8 days), maximum of 31 days and minimum of 11 days.

Clients' Disability Classification

In administering the COPM, the participating therapists were required to classify their clients' disability according to the World Health Organization's classification. Clients could be classified into more than one disability. A summary of the therapists' classifications of the clients' disabilities is shown in Table 4.2. As shown, the distributions of types of disabilities within the orthopedic and stroke groups were similar. "Personal care" and "locomotor" were the two most frequently identified disabilities for the clients in the orthopedic group and among the three most frequently identified disabilities in the stroke group. The finding that personal care and locomotion were frequently identified disabilities is similar to the finding reported by Law et al. (1994) which suggested that personal care and locomotor problems are the two most common domains of concern of occupational therapists. The third disability within the set of most frequently identified disabilities of the stroke group was problems with dexterity. In contrast, only one client in the orthopedic group was classified this way. This difference is attributable to the fact that the pathology of stroke usually affects an entire side of a client's body including the upper and lower limbs, hence compromising hand dexterity. However, pathology of orthopedic problems is localized to a certain bone structure or bodily joint. Common traumas such as fractures of the femur or the hip joint are seldom associated with the dexterity of a client.

Table 4.2

Types of Disability of Clients (COPM)

Disability (WHO Classification)	Percentage (%) Orthopedic $n = 30$	Percentage (%) Stroke $n = 9$
Behavior	0.0	11.1
Communication	0.0	0.0
Personal Care	73.3*	55.6*
Locomotor	80.0*	66.7*
Body Disposition	23.3	33.3
Dexterity	3.3	66.7*
Situational	30.0	0.0
Particular Skills	6.7	22.2
Other Activity Restrictions	30.0	0.0

Note. * denotes over 50.0% of the clients presenting the disability.

COPM Initial and Pre-discharge Assessment Scores

After the therapists had classified the type(s) of disability of their clients, the clients were asked to review their daily activities at home and identify particular activities where they thought they might have problems with. Their answers reflected the daily life habits and activity patterns of the clients.

The clients were asked to rate Importance of those activities that were identified as difficult to perform. Then the clients were asked to rate their Performance and Satisfaction with the five activities with the highest Importance ratings. The five activities selected for the additional ratings are called "activity items". The computations of the Performance and Satisfaction subscale scores of each activity item were as follows:

Performance Score = Importance Rating x Performance Rating

Satisfaction Score = Importance Rating x Satisfaction Rating

The total Performance score of the initial assessment (PERSC 1) is the mean of the Performance scores of all the activity items. The total Satisfaction Score of the initial assessment (SATSC 1) is the mean of the Satisfaction scores of all the activity items. The same is applied to the total Performance Score (PERSC 2) and total Satisfaction Score (SATSC 2) of the pre-discharge assessment. The comparisons of the mean scores on the COPM of the initial and pre-discharge assessments are presented in the following paragraphs. Detailed analyses of the types of activities identified by clients between different diagnostic groups and life roles of clients are presented in the substantive validity section of Chapter V.

The mean and standard deviation of the Performance and Satisfaction subscale scores of the initial and pre-discharge assessments on the COPM are reported in Table 4.3. A two-way (group-by-occasion) repeated measures ANOVA was used to compare the mean subscale scores on the COPM of the orthopedic group with the mean subscale scores on the COPM of the stroke group and at different assessment occasions. As well, the interaction between diagnostic group and assessment occasion was tested. The analyses were carried out separately for the Performance and Satisfaction subscales.

Table 4.3

Subscale Scores on the COPM of Clients

	Initial Ass. Scores				Pre-discharge Ass. Scores			
	PERSC 1		SATSC 1		PERSC 2		SATSC 2	
Diagnostic Groups	M	SD	M	SD	M	SD	M	SD
All (N=39)	25.3	14.7	25.5	17.0	59.5	21.0	59.5	22.9
Orthopedic (*n*=30)	23.0	14.4	23.8	17.0	58.7	21.1	57.8	23.0
Stroke (*n*=9)	32.7	13.8	31.2	16.2	62.3	21.7	65.0	22.8

Although the mean Performance subscale scores of clients in the stroke group were higher than those of the orthopedic group in both the initial and pre-discharge assessments (Table 4.3), there was no significant main effect for diagnostic group on clients' Performance scores ($F=1.64$; $df=1, 37$; $p<.21$). However, there was a significant difference in the Performance scores between the initial and pre-discharge assessments ($F=52.69$; $df=1, 37$; $p<.01$). The interaction between diagnostic groups and assessment occasions was insignificant ($F=.45$; $df=1, 37$; $p<.51$). Similar to the Performance scores, the mean Satisfaction subscale scores of clients in the stroke group were higher than those of the orthopedic on both assessment occasions (Table 4.3), but the main effect for diagnostic groups was not significant ($F=1.47$; $df=1, 37$; $p<.23$). There was a significant difference in clients' mean Satisfaction scores between the initial and pre-discharge assessments ($F=52.30$; $df=1, 37$; $p<.01$). The interaction between diagnostic group and assessment occasion was not significant ($F=.01$; $df=1, 37$; $p<.98$). For both Performance and Satisfaction subscales, the mean scores at the time of pre-discharge were significantly greater than the mean scores at the time of the initial assessment.

The mean Performance and Satisfaction subscale scores for the total sample are similar in value to those reported in Law et al.'s (1994) pilot testing of the COPM. In Law et al., the initial and pre-discharge Performance subscale scores were 34.12 (SD = 15.70) and 63.16 (SD = 19.15) respectively. In the present study, the Performance subscale scores were 25.3 (SD = 14.7) and 59.5 (SD = 21.0). Independent *t*-tests revealed no significant differences in the mean initial Performance subscale scores ($t=.56$, $df=293$, $p>.05$) and the mean pre-discharge scores ($t=.19$, $df=176$, $p>.05$) between the two studies. In Law et al.'s study, the initial and pre-discharge Satisfaction subscale scores were 32.35 (SD = 16.35) and 63.32 (SD = 19.62) respectively, whilst the Satisfaction subscale scores of this study were 25.5 (SD = 17.0) and 59.5 (SD = 22.9). Similarly, no significant differences were found in the mean initial Satisfaction subscale scores ($t=.42$, $df=292$, $p>.05$) and the mean pre-discharge scores ($t=.19$, $df=176$, $p>.05$) between the two studies.

Analyses of clients' results on the COPM suggested no significant differences between the stroke and orthopedic groups. As a consequence, diagnosis was not used as a grouping variable in subsequent COPM related analyses.

COPM Performance and Satisfaction Subscale Scores

The mean Performance and Satisfaction subscale scores, on both assessment occasions, were found to be similar to one another (Table 4.3). Insignificant differences were revealed for both the initial assessment ($t=-.16$, $df=38$, $p<.88$) and the pre-discharge assessment ($t=.04$, $df=38$, $p<.97$). The high correlations between the Performance and Satisfaction subscale scores in the initial assessment

($r=.77$, $p<.01$) and the pre-discharge assessment ($r=.94$, $p<.01$) raised plausible hypotheses which were likely to threaten the validity of the COPM. Are the perceptions of self performance and satisfaction with performance different constructs? Does the COPM measure some other undesirable constructs rather than clients' self perception of performance and clients' satisfaction with their performance? These questions will be investigated more thoroughly in chapters V and VI.

The significant differences of the mean subscale scores on the COPM between the initial and pre-discharge assessments reflected the changes within the clients in both diagnostic groups. This change is attributable to receiving occupational therapy and other rehabilitation programs in the hospital. However, the similarities between the Performance and Satisfaction subscale scores on the COPM did not provide adequate information on the particular attribute(s) of the clients that had changed. Data analyzed at the present stage did not attempt to tackle this issue. In-depth exploration of the content-related, substantive, criterion-related evidence on the COPM's construct validity will shed light on addressing the question: "what are the meanings of the scores on the COPM and how are these scores interpreted?"

COPM Performance and Satisfaction Change Scores

The COPM Test Manual suggests that clinicians use differences in scores between the test and retest, that is, the Change in Performance and Change in Satisfaction by subtracting the mean subscale scores of the initial assessment from those of the pre-discharge assessment. This strategy of using mean differences in scores is not recommended by measurement and test specialists (Crocker & Algina, 1986) because it amplifies errors of measurement associated with the final difference scores. Nevertheless, the results of Changes in Performance and Satisfaction scores of clients are presented in Table 4.4. The Change in Performance and Satisfaction scores revealed in this study were 34.3 (SD = 22.8) and 34.0 (SD = 24.4), respectively. In Law et al.'s (1994) study, the corresponding Change scores were slightly lower: the mean Performance change was 27.78 (SD = 17.33) and the mean Satisfaction change was 29.4 (SD = 19.90). No comparisons of the difference scores were made between the Performance and Satisfaction subscales within the two diagnostic groups. It is important to note, however, the large standard deviations of the difference scores. This will play an important role in estimating their reliabilities and standard errors of measurement.

Table 4.4

Mean Performance and Satisfaction Change Scores of Clients

Diagnostic Groups	Change in Performance		Change in Satisfaction	
	M	SD	M	SD
All	34.3	22.8	34.0	24.4
Orthopedic	35.7	22.8	34.0	25.0
Stroke	29.7	23.5	33.8	23.3

Results of Satisfaction with Performance Scaled Questionnaire

The SPSQ required the clients to rate their satisfaction on 46 items categorized into two sections. The first subscale assesses the degree to which clients are satisfied with their performance in home management tasks and the second subscale assesses social and communication problem solving tasks. A higher score in each subscale reflects a higher degree of satisfaction.

Table 4.5

Mean SPSQ Home Management Scores (HMS) of Clients

	Initial HMS		Pre-discharge HMS	
Diagnostic Groups	M	SD	M	SD
All	61.7	37.1	69.2	35.4
Orthopedic	59.0	38.9	65.4	37.1
Stroke	70.4	30.4	82.0	27.4

As with the COPM, the means of the Home Management scores of the two diagnostic groups and at the two assessment occasions were compared using two-way repeated measures ANOVA. In the initial assessment, the mean HMS on the SPSQ of clients in the stroke group was higher than that of the orthopedic group (Table 4.5). The same pattern was found in the pre-discharge mean HMSs. Higher mean HMSs were shown in the pre-discharge assessment when compared with the initial assessment results. However, neither the main effects for diagnostic group nor assessment occasion was significant (Group: $F=1.19$; $df=1, 37$; $p<.28$; and Occasion: $F=3.23$; $df=1, 37$; $p<.08$). The interaction effect was also insignificant ($F=.28$; $df=1, 37$; $p<.60$).

The SPSQ mean Social/Communication scores showed the reverse pattern when compared with the Home Management subscale (Table 4.6). The orthopedic group had a higher mean SCSs than the stroke group. The mean SCSs of clients in the pre-discharge assessment were higher than those in the initial assessment. However, results of two-way repeated measures showed no significant effects (Groups: $F=1.93$; $df=1, 37$; $p<.17$; Occasion: $F=3.45$; $df=1, 37$; $p<.07$; and Interaction: $F=1.40$; $df=1, 37$; $p<.25$).

Table 4.6

Mean SPSQ Social/Community Problem Solving Scores (SCS) of Clients

	Initial SCS		Pre-discharge SCS	
Diagnostic Groups	M	SD	M	SD
All	53.8	24.0	58.6	20.8
Orthopedic	57.3	26.0	59.9	20.8
Stroke	42.2	9.56	54.2	21.0

The failure to find statistical significance may be attributable to the insensitivity of the SPSQ or the limited number of clients. This phenomenon could impede the quality of the evidence on the convergent relationships between the

SPSQ and the Satisfaction subscale of the COPM and the divergent relationships between the SPSQ and the Performance subscale of the COPM. Findings on these relationships will be discussed in Chapter V.

Results of Klein-Bell ADL Scale

The Klein-Bell ADL Scale (KBADL) requires the participating therapists to assess and rate their clients' performance in self-care activities according to a standardized format and rating scale. There are 170 items divided into six areas of self-care task performance. The total score reflects clients' competence and independence in performing various self-care activities. A higher KBADL score means a higher degree of independence in self-care functioning.

In general, clients in the orthopedic group had higher means scores on KBADL than those in the stroke group in both the initial and pre-discharge assessments (Table 4.7). Results also suggested that the mean KBADL scores in the pre-discharge assessment were higher than those of the initial assessment for both diagnostic groups. Two-way repeated measures ANOVA revealed insignificant differences in the mean KBADL scores of clients between the orthopedic and stroke groups ($F=2.60$; $df=1, 37$; $p<.12$). However, significant differences were found in the mean KBADL scores of clients between the two assessment occasions ($F=67.68$; $df=1, 37$; $p<.01$). The interaction between the two diagnostic groups and the two assessment occasions was not significant ($F=.94$; $df=1, 37$; $p<.34$).

Table 4.7

Mean Klein-Bell ADL Scale Scores of Clients

	Initial KBADL		Pre-discharge KBADL	
Diagnostic Groups	M	SD	M	SD
All	233.5	35.6	274.7	24.2
Orthopedic	238.5	30.8	277.2	18.2
Stroke	217.1	46.7	266.2	38.2

Results of the Klein-Bell ADL Scale suggested that clients' performance in self-care activities measured by the instrument had improved throughout the course of occupational therapy and other rehabilitation programs in the hospital. The initial and pre-discharge assessments captured the changes within the clients. This phenomenon was consistent with the trends revealed by the COPM which also showed significant changes in the clients' perceived performance. However, there were two major differences between the KBADL and the COPM. First, the KBADL used an expert rating whereas the COPM used a self-rating method. Second, the KBADL used a pre-determined domain of self-care activities against which clients' performance were rated. In contrast, the COPM adopted a client-centered domain with the possibility that the activities selected by each client could be totally different from those selected by other clients. The results of the KBADL will be used to establish evidence convergently correlated with the results of the COPM in Chapter V.

Results of Functional Independence Measure

The FIM uses an expert-rated performance assessment format. Clients are rated on 18 items: 13 in the motor subscale and 5 in the cognitive subscale. The mean FIM motor and cognitive scores reflect the constructs of functional status, severity of disability, and burden of care. Higher scores on the FIM subscales reflect higher general functional status, and, consequently the less severity of disability and burden of care of the clients.

In general, clients in the orthopedic group had higher mean FIM Motor scores than those in the stroke group in both the initial and pre-discharge assessment occasions (Table 4.8). Results also showed that the pre-discharge Motor scores were higher than those of the initial assessment in both diagnostic groups. Two-way repeated measure ANOVA revealed significant differences between the mean Motor scores of the two diagnostic groups (F=8.59; df=1, 37; $p<.01$) and the two assessment occasions (F=56.11; df=1, 37; $p<.01$). However, the interaction between diagnostic group and assessment occasion was not significant (F=1.30; df=1, 37; $p<.26$).

Table 4.8

Mean FIM Motor Scores of Clients

	Initial FIM Motor Score		Pre-discharge FIM Motor Score	
Diagnostic Groups	M	SD	M	SD
All	66.3	13.9	80.8	7.6
Orthopedic	69.2	10.8	82.5	3.6
Stroke	56.9	19.2	75.0	13.3

Results of the FIM Cognitive subscale scores showed smaller standard deviations than their Motor subscale counterparts, especially the pre-discharge Cognitive scores (Table 4.9). These were likely to have been attributable to the ceiling effect in the pre-discharge assessment since the maximum Cognitive subscale score is 35. Significant differences were revealed between the mean Cognitive subscale scores of the orthopedic and stroke groups (F=7.46; df=1, 37; p<.01). The difference between the Cognitive subscale scores obtained on the two assessment occasions was also statistically significant (F=15.39; df=1, 37; p<.01). In addition, the interaction between diagnostic groups and assessment occasion was significant (F=5.63; df=1, 37; p<.02). This is consistent with the hypothesis that the significantly higher pre-discharge Cognitive subscale score of clients in the stroke group was due to the combined effects of the cognitive dysfunction of clients suffering from stroke and the effects of the treatment programs that they received.

Table 4.9

Mean FIM Cognitive Scores of Clients

	Initial FIM Cognitive Scores		Pre-dis. FIM Cognitive Score	
Diagnostic Groups	M	SD	M	SD
All	32.9	3.7	34.2	1.6
Orthopedic	33.7	3.3	34.5	1.4
Stroke	30.1	4.1	33.2	1.9

Results of the Motor and Cognitive subscales of the FIM suggested that clients' performance in selected motor and cognitive tasks improved from the initial to the pre-discharge assessment occasions. The FIM was used to investigate the convergent relationships of the COPM.

Conclusion

Results of the clients on the COPM and other clinical instruments were presented in this chapter. Significant increases were observed in the mean scores of the pre-discharge assessment when compared with those scores of the initial assessment. This observation was true for all the clinical instruments with the exception of the SPSQ Home Management subscale. This suggests that clients in general experienced improvement in both their performance and satisfaction with their own performance. No significant differences, however, were found between the assessment results of the clients in the orthopedic and stroke group except for the Motor and Cognitive subscale scores of the FIM.

CHAPTER V

EVIDENCE OF CONSTRUCT VALIDITY

Introduction

The evidence concerning the construct validity of the COPM is presented in this chapter. The basic framework of investigation and the presentation of results are based on the validation models suggested by Messick (1989) and Loevinger (1957). This chapter is divided into four sections, with each section describing one type of evidence. These sections are entitled content-related evidence, substantive evidence, structural-related evidence, and criterion-related evidence. In each of these sections, the methods with which evidence was gathered and the results of data collection are described in detail. At the end of each section, a conclusion is drawn to summarize the discussion and its implications. Integration and suggestions for changes and improvement of the COPM are presented in Chapter VII.

Section A - Content-Related Evidence

In Section A of this chapter, the content-related evidence contributing to the construct validity of COPM is presented. Data gathered from the expert panel review were used to assess the relevance and representativeness of the content of COPM. As noted in chapter III, nine panel members were recruited to evaluate the content, structure, and clinical utility of the COPM using a 38 item questionnaire. The members spent on average one hour to complete the questionnaire. The content-related evidence reported in this section is based on the 11 item Content Review section of the questionnaire. The findings of the Structural Review are discussed in Section C of this chapter; and the results of the Utility and Procedural Review are presented in chapter VI.

The Content Review Section

This section was designed to gather the opinions of the panel members about the relevance and representativeness of the content (test domain) of the COPM with reference to the model of occupational performance upon which the test was built. An example of the items used in this section is:

1.2 How well does the COPM reflect the productivity performance of a person being assessed?

1	2	3	4	5
Poor	Fair	Good	Very Good	Excellent

Which part of the COPM do you think assesses the productivity performance?

The above item attempted to evaluate the relevance of the COPM as a measure of a client's performance in the area of productivity (one of the three areas of occupational performance). Members were required to make responses using the five point scale (1 - poor and 5 - excellent). These responses were then justified in the open-end portion.

The first nine items (1.1 - 1.9) assessed the relevance of the COPM's test content to the specific components of the model of occupational performance (relevance). The next two items, 1.10 and 1.11, were summative items. Item 1.10 measured the extent to which the test construct of COPM represented the construct of occupational performance (representativeness). Item 1.11 measured an overall relevance of the test content to the theoretical model.

Relevance of COPM's Test Content to Evaluation of Occupational Performance

The first four items (1.1 to 1.4) in the Content Review section asked about the relevance of the COPM's content and assessment processes in evaluating clients' occupational performance. Items 1.1 to 1.3 covered the three areas of human performance, namely self-care, productivity, and leisure. Item 1.4 measured an overall relevance of the test content to the assessment of clients' occupational performance. Panel members' mean rating in this domain was 3.0. This suggests that the perception of the relevance of the COPM to the assessment of occupational performance was "good" (Table 5.1). The same level of relevance was also applied to the three areas of occupational performance, i.e., self-care, productivity, and leisure. However, two raters, Raters 1 and 2, regarded the COPM as "not relevant" (i.e., ratings of 1 or 2) to the assessment of productivity and leisure activities.

Table 5.1

Relevance of COPM to Assessment of Occupational Performance

Item No.	Raters									Item Mean Scores
	1	2	3	4	5	6	7	8	9	
1.1 Self-Care Performance	3	3	5	3	3	3	3	3	3	3.2
1.2 Productivity Performance	2	1	4	4	3	3	3	3	3	2.9
1.3 Leisure Performance	2	1	3	4	3	3	3	3	3	2.8
1.4 Overall Relevance	2	2	4	3	3	3	4	4	3	3.1
	2.3	1.8	4.0	3.5	3.0	3.0	3.3	3.3	3.0	3.0

To gain more insight into the panel members' evaluation of the relevance of COPM to the assessment of occupational performance, the responses made by the panel members in the open-ended portions of the four items were analyzed. Responses which identified the particular content and process of the COPM as relevant to the assessment of occupational performance formed one group.

Comments which identified parts of the COPM as irrelevant to the assessment formed the second group.

Six out of the seven panel members who provided written responses commented that the assessment of clients' occupational performance occured in the processes of the semi-structured interview and Problem Definition (Step 1) in Self-Care (Step 1a), Productivity (Step 1b), and Leisure (Step 1c). Moreover, the processes of clients' self-rating of Importance, Performance, and Satisfaction were also perceived as relevant to the assessment of occupational performance.

However, four of the members criticized the accuracy and adequacy of this format of assessment. They pointed out that the results obtained were largely influenced by the extent to which the clients understood the processes and had insights into their own situations and what problems would exist after their discharge. Another panel member felt that clients' rating of performance and satisfaction relied very much on whether the clients had the experience of trying out the activities which they selected for rating. These flaws contribute to irrelevant test variance (Messick, 1993, p.34) which confounds interpretation of the information obtained from the COPM. One panel member commented,

> "The ability of clients to determine their self-care areas sets precedence to how well the COPM measures self-care. If a patient is able to do this (identifying problems), then the performance and satisfaction scores can act as a measure....the ability of the persons was assessed to how well they

> identify problems, and determine their performance and satisfaction with it."

Two other members wrote,

> "Whether or not the items are adequate depends on clients' insight."
>
> "It is difficult for patients to rate activities if they have not actually tried them out."

A fourth member further substantiated this point of view,

> "They (clients) may have considerable difficulties (performing particular activities) but not be able to understand the implication of their injury on their ability."

Comments of the panel members revealed that majority of them agreed with the relevance of the content of the COPM to the assessment of clients' occupational performance. Comments from panel members reflected that the identification of perceived problems and ratings of performance and satisfaction with performance demanded clients' abilities to understand the implications of their own disabilities and environment after their discharge from hospital. In other words, clients were required to have intact higher cortical functioning in order to give valid responses to the COPM. Panel members' criticisms highlighted another source of "irrelevant variance" in the COPM test scores which further confounds interpretation of the test results. For a client who identified only one problem with performance, it is difficult to differentiate whether that particular client was able to manage most of the performances or whether the client was merely making an incompetent decision due to the lack of insight. Similarly, high scores on the Performance and Satisfaction subscales do not necessarily reflect good performance and satisfaction with performance of the client. Instead, it could be a reflection of poor understanding of the impacts of disability of the client. The COPM does not have the mechanism to differentiate the confounding factors from its assessment of occupational performance, such as an item validity check or screening component.

Relevance of COPM to Assessment of Performance Components

According to the construct of occupational performance, mental, physical, sociocultural, and spiritual components are basic functions of each individual from which activities can be carried out. Performance components can involve objective measurements of individuals' abilities and life situations, such as emotions, cognitive abilities, neuromotor abilities, social network, home environment, beliefs, and attitudes. Individuals are characterized by their own performance components. Individuals' performances in the areas of self-care, productivity, and leisure are the products of interactions between their performance components and their unique environment and life roles.

The previous section has shown that the relevance of the content and testing processes of the COPM to the model of occupational performance was rated overall as good. Table 5.2 summarizes the results of the relevance of the COPM as a measure of the four basic building blocks of an individual's functioning relative to the performance components. These building blocks are the mental, physical, sociocultural, and spiritual components.

Table 5.2

Relevance of COPM to Assessment of Performance Components

Item No.	Raters									Item Mean Scores
	1	2	3	4	5	6	7	8	9	
1.5 Mental	1	2	1	1	2	2	2	2	1	1.6
1.6 Physical	2	2	2	4	3	3	1	3	1	2.3
1.7 Sociocultural	3	3	1	3	3	2	2	4	1	2.4
1.8 Spiritual	2	1	1	2	2	2	3	4	1	2.0
1.9 Overall Relevance	2	2	2	3	3	3	2	3	1	2.3
Mean Scores	2.0	2.0	1.4	2.6	2.6	2.4	2.0	3.2	1.0	2.1

Five items (1.5 to 1.9) were constructed to gather panel members' opinions on the relevance of the COPM to the assessment of the performance components with reference to the construct of occupational performance. Items 1.5 to 1.8 covered the mental, physical, sociocultural, and spiritual performance components.

Item 1.9 was to evaluate the overall relevance of the test content to the assessment of clients' performance components. The mean rating for this domain was 2.1 suggesting that the COPM provided a "fair" measure of the clients' performance components (Table 5.2). The results of the relevance review were presented separately for each performance component.

Mental Performance Component. The mean rating of the relevance of using the COPM to assess mental functioning was 1.6 (Table 5.2); the ratings across panel members were either 1 or 2. The low mean and the consistency of panel members' ratings reflect the opinion that the COPM is not relevant to assessing mental functioning. Four panel members commented that mental functioning was not directly and objectively assessed. Rather, it was indirectly evaluated by observing clients' abilities to follow instructions, make a list of activities, and use the rating scales.

According to the model of occupational performance, the mental performance component is the total emotional and intellectual abilities of an individual (Appendix I). Clients' responses on the COPM were not found to directly nor objectively reveal their mental functioning. The identification of problems and the Performance and Satisfaction subscale scores did not allow inferences to be made on clients' emotional status, memory, or comprehension skills. As two of the panel members commented,

> "...none reflect this (mental performance component). This is picked up by the clinician as an 'aside'."
>
> "I don't think the test measures this (mental performance component). Probably talking with the clients gives you an idea of their emotional and intellectual functioning."

Two panel members further commented that the reasons with the COPM failed to measure the clients' mental performance component were the subjectivity involved and the lack of directions for documenting clients' deficits:

> "When a client is asked to list the activities and weight those problems, the measure is not objective."
>
> "The interview and rapport with clients help determine their mental performance component....If the clients are unable to detect problem areas or if they are very vague, it is also a reflection of their mental performance. Nowhere to document mental performance."

To summarize, the panel members unanimously agreed that the content of the COPM was not a direct and adequate measure of clients' mental performance component. However, the processes of involving clients to identify their problems and rate their performance and satisfaction with performance provided opportunities for the clinicians to observe and indirectly assess the mental functioning of clients. However, as mentioned in the previous section, these mental functions were mainly at a higher cortical level such as understanding of and insights into their disability and associated problems. The indirect assessment of them as in the COPM is far from adequate. Inferences made on clients' mental performance component by interpreting COPM's test results are therefore not meaningful. Further, neither the method of documentation, nor the objectivity required in the assessment of this performance component are provided in the COPM and this omission was identified as a weakness in the instrument. Incorporation of the assessment of these aspects may be indicated in future revision of the instrument.

Physical Performance Component. The mean rating of the relevance of the COPM to the assessment of the physical performance component was 2.3 which falls between "fair" and "good" (Table 5.2). This score suggests that the COPM was found to better reflect physical functioning of clients than it did mental functioning.

Two panel members pointed out that assessment of physical functioning with the COPM only occured when clients' problems were attributable exclusively to their physical dysfunction. For example, a client with a stroke may have problems shopping in a mall due to poor motor control and walking balance. Another client may have the same problem with shopping but for a different cause - perhaps inadequate practice or lack of self-confidence in performing such an activity. These two scenarios suggest that clients may present with exactly the same dysfunction in their occupational performance, but the causes of their dysfunction, involving physical and/or mental performance components, may be widely different. Nevertheless, one member who gave a "Very Good" rating wrote the following comment:

> "All steps of identifying problems focus on motor ability of clients, and scoring the performance is mainly based on motor or physical skills."

This opinion suggests that the COPM is an adequate measure of the physical performance component assuming that performance of self-care, productivity, and leisure activities are largely determined by physical capability.

In summary, the relevance of content and processes of the COPM as a measure of the physical performance component was rated as less than good. Results of the COPM may directly reflect clients' problems in the physical component, but only if the dysfunction originated solely from a physical condition, such as muscle weakness, poor endurance, or sensory deficits. Incorporation of standardized procedures to assess clients' physical performance component may be indicated in future revision of the instrument.

The design of the COPM makes the instrument unable to isolate dysfunction attributed to mental and physical performance components and impossible to be interpreted at a performance component level. Instead, valid clinical inferences rely heavily on the interpretation made by the clinician who administers the test. These clinicians need also to have substantial clinical experience in particular areas, such as orthopedic or stroke.

Sociocultural Performance Component. The mean rating on the relevance of the COPM to the assessment of the sociocultural performance component was 2.4 which means between "fair" and "good". The variation of ratings among the panel members may be attributable to their different beliefs in objectivity versus subjectivity in clinical evaluation. Their comments, however, were generally consistent. Three of the members indicated that the sociocultural performance component was indirectly reflected by the list of activities that the clients selected and the Importance rating given to those activities. This is illustrated by the following comments:

> "Each problem listed and weighted is a reflection of the client's value system and life situation."
>
> "Identification of problem areas and rating of importance to overcome the problems (e.g., returning to work) can reflect client's beliefs and values in his or her life."

Unlike the mental performance component, the COPM reveals clients' sociocultural component such as their valuing of various activities (Importance ratings), life roles (as discussed in Chapter IV), and life situation. Similar to the two performance components mentioned previously, the COPM does not give a distinct account of a client's sociocultural performance component. The inference from interpreting the results of the COPM about the clients' sociocultural performance component would only be appropriate if the problems of the clients are attributable solely to their dysfunction in the sociocultural aspects. Incorporation of standardized procedures to assess clients' sociocultural performance component may be indicated in the future revision of the instrument.

Spiritual Performance Component. The mean rating on the relevance of the COPM to the assessment of the spiritual performance component was 2.0 (Table 5.2). This rating suggests that relevance of the content and processes of the COPM as a measure of spiritual performance component is fair. The comments from the panel members indicated that the COPM indirectly assessed the spiritual performance component. Further, this indirect evaluation came from clients' rating on the "Satisfaction" with performance subscale. Two of the members pointed out that the ratings on the Importance subscale could also reflect spiritual performance. One panel member wrote,

> "It is difficult to distinguish (the spiritual component) from the sociocultural aspect."

This member also stated that the spiritual component was assessed:

> "...partially in the Satisfaction and Importance scores."

Another panel member added this critique:

> "The COPM could be improved in [the] area of ... labelling what is important for them (clients) to work on and become better at to improve their quality of life. It does not really probe into these depth (refer to human being's spirituality), but only indirectly."

The content and testing processes of the COPM did not provide a direct measure of spiritual performance component of clients. Incorporation of standardized procedures to assess clients' spiritual functioning may be indicated in the future version of the instrument.

Summary. The mean rating on Item 1.9 was 2.1 indicating that panel members' overall evaluation of relevance of the content and processes of the COPM as a measure of the four performance components was fair. This overall rating was consistent with the mean ratings on individual components reported previously (ranged from 1.6 to 2.4). A common criticism of all areas was that the interpretation of the results of the COPM failed to provide separate information on clients' competence in each performance component. Inferences on clients' performance components based on the COPM's test results can only be made indirectly either by the assessing clinicians or by the use of supplementary clinical evaluation.

Overall Relevance & Representativeness to Construct of Occupational Performance

As mentioned earlier in this chapter, the nine items in the questionnaire cover the relevance of the COPM in terms of assessing occupational performance (three areas) and performance components (four factors). Item 1.10 in the questionnaire was constructed to gather panel members' opinions on the overall notion of representativeness.

Previous discussion in this section revealed that the COPM was regarded as a "good" measure of clients' occupational performance, i.e. of doing particular activities. However, content and testing processes were given only "fair" ratings in the ability of the COPM to capture clients' dysfunction in their performance components. Since the model of occupational performance includes two levels of characteristics of individuals, this suggests that the COPM is likely not able to completely represent the model when it is used as a clinical measurement tool to make inferences on clients' occupational performance.

Table 5.3 summarizes panel members' rating on the COPM's representativeness of the construct of occupational performance (item 1.10). The mean rating of 2.4 suggests that members had reservations about the representativeness of the COPM; five panel members rated it as fair while four rated it as good in the match between the construct of occupational performance to the test domain.

Table 5.3

Overall Relevance and Representativenss of COPM to Model of Occupational Performance

Item No. Overall Model of Occupational Performance	Raters									Item Mean Scores
	1	2	3	4	5	6	7	8	9	
1.10 Complete Represent.	2	2	3	3	3	2	2	3	2	2.4
1.11 Overall Construct Relevance	3	4	4	4	5	4	4	5	3	4.0
Rater Mean Scores	2.5	3.0	3.5	3.5	4.0	3.0	3.0	4.0	2.5	3.2

Comments from panel members were consistent in reflecting problems and weaknesses with the use of COPM as an instrument assessing clients' occupational performance. First, the COPM is inadequate in the assessment of performance components, especially the mental and spiritual aspects. Although clients' problems with occupational performance are supposed to be the outcomes of the dysfunction in performance components, the assessment at the activity level does not seem able to provide enough insights for clinicians to make inferences, plan and implement appropriate interventions. As one of the panel members wrote,

> "The COPM represents clients' occupational performance, however, it does not adequately measure the performance components to develop a treatment plan (although occupational performance is supposed to reflect the components) for a client."

Second, the COPM does not formalize the observations and inferences that clinicians want to make during the testing processes so that they can have a better understanding of clients' problems in occupational performance. For example,

> "The COPM identifies the areas of self-care, productivity, and leisure. The COPM, in my opinion, does fall short in the areas of mental and spiritual dysfunction which are not directly measured. Clinicians may be able to determine the dysfunction through [the] interview process and it is very much relied on the on-site performance of the client."

Third, the interview processes used with clients when administering the COPM are critical in eliciting relevant information and building therapeutic interactions. However, these processes are not standardized, thereby leading to substantial variations in the adequacy, quality, and specificity of the information obtained. Another panel member wrote,

> "It is good to measure physical ability. Ranking problems and scoring them bring in client's concerns and perception [sic]. However, most of the focus remains on physical ability."

The last item (1.11) in the Content Review is a summative item constructed to obtain a global impression of the relevance of the COPM to reflect clients' functioning with respect to the construct of occupational performance. Surprisingly, the mean rating on this item was 4.0 suggesting that results obtained from the COPM were thought to be "very good" in their relevance and appropriateness. The individual opinions ranged from "good" to "excellent". This seems at odds with the ratings given for the individual performance components. Two members who gave an overall rating of "excellent" seemed to share the opinion that the processes involved in the COPM were extremely relevant to the ideology of the client-centred approach in the assessment of clients' occupational performance. One member wrote:

> "This is very relevant because it provides the opportunity for the client to take some control to indicate what they view as important and how they perceive themselves to be doing which is then incorporated into therapy regime to empower the client and his rehabilitation."

It is clear from this comment that in (at least) one case, the excellent rating actually referred to the processes that the clients went through in completing the COPM rather than on the test content of the instrument. In fact, most of the members agreed that the client-centred approach had been operationalized in the COPM.

One weakness of the COPM seems to be that the testing process does "not work well with a client with decreased insight and cognitive problems". This combined with the other shortcomings mentioned in the previous sections indicates that there are many problems in the content-related validity of using the COPM to make meaningful and truthful inferences of clients' occupational performance.

Section B - Substantive Related Evidence

Two different strategies were used to analyze the test content and the testing processes the clients engaged in when completing the assessment. The first strategy was to relate the activities the clients perceived as being difficult and of concern in the COPM with their responses to the Role Checklist. The second strategy was to analyze the test taking processes of the clients as they were assessed by the COPM. Protocols were obtained through semi-structured interviews in which clients were asked to describe the thought processes that they used when they responded to the tasks of problem identification, importance rating, perceived performance, and satisfaction with performance. The clients were also asked to elaborate on the responses that they gave in completing the COPM.

Response Consistency in Process of Problem Identification

According to Messick (1989) and Loevinger (1957), the substantive component of validity entails looking for consistencies in test content with test responses. In this portion of the validation study, the responses made by the clients in the COPM were analyzed for their convergence with the construct of occupational performance and other relevant constructs. First, the activities identified by the clients were classified using the model of occupational performance. Then a comparison was made of importance ratings and patterns of the identified activities of clients in different gender and diagnostic groups.

Patterns of Activities Revealed in the COPM

At the beginning of the COPM, clients are asked to consider their present abilities and performance of daily living tasks and identify any activity that they perceive as difficult to perform. In the present study, all the activities identified by the clients were first assigned to the activity category list provided in the COPM Test Manual (Law et al., 1991, p.31-34). The list adopts the occupational performance model which divides 99 daily activities into three clusters, namely self-care, productivity, and leisure. Under each cluster of occupational performance, there are three different categories of activities forming a total of nine categories. For example, the first category under self-care occupational performance is personal care which includes 13 activities such as "getting clothes from closets and drawers" and "washing self". According to the authors of the Test Manual, the list only serves as an example for reference and hence it is not exhaustive. For this study, nine more activities were added to the original list to accommodate all of the activities listed by the clients. The revised activities list and their coding as used in this study are presented in Appendix XI.

Results from the COPM showed that clients identified 327 activities which they regarded as difficult to perform in the initial assessment. As shown in Table 5.4, the distribution of the activities with respect to the three areas of occupational performance were: 56.0% (self-care), 22.0% (productivity), and 22.0% (leisure). These results are similar to those reported by Law et al. (1994): 54.2% (self-care), 25.6% (productivity), and 20.1% (leisure). The patterns of activities identified by the clients in these two studies suggest that in the initial assessment most of the clients were concerned about their independent performance in self-care activities.

Among the self-care activities, clients were found to value functional mobility (26.3%), such as walking and getting in/out of a tub, and personal care (19.3%), such as dressing and washing self, more than community management (Table 5.4). Clients also showed more concern for their performance of home management (16.5%) and active recreation (10.4%) than other types of activities under productivity and leisure occupational performance.

Table 5.4

Frequencies of Activity Categories in the COPM Identified by the Clients

Activity Categories	Frequency	Percentage
Self-Care -Personal Care	63	19.3*
Self-Care - Functional Mobility	86	26.3*
Self-Care - Community Management	34	10.4*
Self-Care - Total	183	56.0
Productivity - Paid/Unpaid Work	16	5.0
Productivity - Household Management	54	16.5*
Productivity - Play/School	2	0.6
Productivity - Total	72	22.0
Leisure - Quiet Recreation	22	6.7
Leisure - Active Recreation	34	10.4*
Leisure - Socialization	16	4.9
Leisure - Total	72	22.0
Total	327	

Note. N = 327. (*) denotes the activity categories with 10% or more frequency out of the total number of activities identified by the clients in the COPM.

Listed in Table 5.5 are the 11 activities that were most frequently identified by the clients as being difficult to perform. The frequencies ranged from a high of 21 (6.5%) to a low of 11 (3.4%) out of a total number of 327 activities. The clients showed more concern with activities related to self-care and productivity, than to leisure. In self-care, functional mobility was found to receive the most emphasis. More clients identified walking; getting in/out of cars, tub, and bed; and climbing stairs as difficult to perform in the initial assessment. Personal care, especially in dressing upper and lower limb garments and washing self, was another area of concern identified by the clients. Community management received comparatively less emphasis. The clients selected shopping and driving a car as the two most problematic activities in the area. In productivity, the activities of household management and paid/unpaid work were of greater concern to clients than play/school activities. In performing household management activities, clients showed the most concern with doing laundry. For paid/unpaid work, clients identified working expected hours as important.

Table 5.5

Eleven Activities Most Frequently Identified as a "Concern" in the COPM

Occupational Performance and Activity Categories	Specific Activities	Frequencies (N=327)
Self-Care - Personal Care	Dressing U/E and L/E Garment	19
	Washing Self	15
Self-Care - Functional Mobility	Walking	21
	Getting in/out of Cars	18
	Getting in/out of Tub	14
	Getting in/out of Bed	12
	Climbing Stairs	12
Self-Care - Community Man.*	Shopping	17
	Driving a Car	11
Productivity - Paid/Unpaid Work	Working Expected Hours	11
Productivity - Household Man*	Doing Laundry	14
Total		164

Note. U/E and L/E is upper and lower extremity. Community Man.* is Community Management. Household Man.* is Household Management.

The patterns of activity on the COPM found in the present study matched very well with what was happening when clients are discharged to home or other placements. In these situations, self-care activities such as personal care and functional mobility are basic to their day-to-day functioning. In contrast, community and household management are activities that the clients usually wait for assistance from health care professionals such as home care services, family members, or friends. It was surprising to find that, in the present study, paid and unpaid work (5.0%) was frequently selected by the clients. This indicates that despite their physical disabilities, some of the clients were active in the area of productivity. Although leisure activities were not included in the list, this did not mean that this aspect was not of a concern to the clients. As reported previously in Table 5.4, 22% of the responses made by the clients covered the leisure area of occupational performance. However, the number of clients selecting any particular activity was evenly spread among the activities grouped under quiet recreation, active recreation, and socialization categories. This further reiterates the notion that the needs of clients in all three areas of occupational performance should not be overlooked by occupational therapists.

Besides the 11 activities specified in Table 5.5, there were numerous other activities that were perceived as difficult to perform at the time of initial assessment. The frequencies of each of these activities are presented in the complete list of activities found in Appendix XI.

According to the COPM, clients are asked to select activities only in the initial assessment. In the pre-discharge assessment, clients rate their performance and satisfaction on the same pre-selected activities. Hence no data are available regarding the concerns of the clients at the pre-discharge stage. It is plausible that the clients might have changed their concerns, and hence their choice of activities, if their functional levels and social situation had been altered. Indeed, some clinicians have pointed out this phenomenon. However, the scope of this study did not include this aspect which could pose some threats to the construct validity of the COPM.

Activities and Their Importance

The review of activities selected by clients on the COPM reflects clients' concerns with independent daily living and, to a certain extent, their lifestyles after discharge from the hospital. The importance of these activities to the clients should be consistent with the activities that are performed in particular life roles and lifestyles. Table 5.6 summarizes the mean ratings of the degree of Importance made by the clients in the COPM for the 11 most frequently identified problems with the use the Importance rating scale 1 through 10. A rating of 1 indicated "Not Important At All" and a rating of 10 indicated "Extremely Important".

Table 5.6

Clients' Importance Ratings of the Activities Most Frequently Identified in COPM

Occupational Performance and Activity Categories	Specific Activities	Mean Importance Ratings
Self-Care - Personal Care	Dressing U/E and L/E Garment	9.2
	Washing Self	9.0
Self-Care - Functional Mobility	Walking	9.1
	Getting in/out of Cars	8.3
	Getting in/out of Tub	8.2
	Getting in/out of Bed	9.0
	Climbing Stairs	8.3
Self-Care - Community Man.*	Shopping	6.7
	Driving a Car	7.5
Productivity - Paid/Unpaid Work	Working Expected Hours	8.7
Productivity - Household Man*	Doing Laundry	6.1

Note. Community Man.* represented Community Management.

Household Man.* represented Household Management.

The results revealed that the four mean Importance ratings of 9.0 and above were clustered in the occupational performance of self-care, especially personal care and functional mobility. It seems that when clients perceived their difficulties in performing activities in personal care and functional mobility, they also felt that those activities were very important to their functioning. Three activities in functional mobility, however, were rated as slightly less important (8.2-8.3). These were getting in/out of cars, climbing stairs, and getting in/out of tub. This may be explained by the fact that these activities either can be performed with adaptations to the environment and the use of adaptive equipment, or they are not required to be performed everyday. By the same token, activities under the community management category were deemed less important than those of personal care and functional mobility.

Surprisingly, "working expected hours" under the productivity occupational performance was perceived as important with a mean Importance rating of 8.7. This indicates that the clients in this study who rated "working expected hours" as difficult also perceived the activity as very important to their living. In contrast, the clients who selected "doing laundry" under the category of "household management" did not perceive that the activity was important to them despite its physically demanding nature. The same explanations given for the activities under community management can be given for "doing laundry" and "getting in/out of car". These activities do not need to be performed everyday, and probably helpers can easily perform them for the clients once in a while.

Gender, Diagnostic Group, and Patterns of Activity

In order to have a more in-depth understanding of the patterns of activities of concern to the clients, the results on the COPM Problem Definition were further analyzed by using diagnosis and gender as grouping variables. When the frequencies of each activity category selected by the clients were grouped under the clients' gender, no significant differences were found in the patterns of activity selection between the male and female clients (*Chi-Square*=6.17, *df*=8, $p > .05$). This means that the patterns of activity of concern to male and female clients in the initial assessment were similar. In contrast, significant differences were found between the activities selected by clients in the orthopedic and stroke groups (*Chi-Square*=28.13, *df*=8, $p < .01$).

Further examination revealed that 30.3% of the activities identified by clients in the orthopedic group were related to functional mobility compared to 15.7% by the clients in the stroke group (Table 5.7). In contrast, clients in the stroke group put more emphasis on whether they could perform activities related to quiet recreation (14.6% of the total responses) than did the subjects in the orthopedic group (3.8% of the total responses). Clients in both groups placed approximately equal emphasis on whether they could perform personal care and household management tasks.

Discussion in this section provides insight into the concerns of the clients based on their responses to the Problem Identification part of the COPM. It is also a preliminary analysis on the meaningfulness of the activities identified by the clients before establishing their relationships with the life roles of the clients. To summarize, information on the types of activities selected by the clients in the COPM reflect particular activity patterns and concerns with their different areas of occupational performance. Findings suggests that there are differences between the clients in the orthopedic and stroke groups. However, these results should be interpreted with caution because of the small sample sizes.

Table 5.7

Comparisons of Frequencies of Type of Activities Selected by Orthopedic and Stroke Groups

	Orthopedic Group (*n*=30)		Stroke Group (*n*=9)	
Activity Categories	*F*	%	*F*	%
Self-Care -Personal Care	45	18.9%*	18	20.2%*
Self-Care - Functional Mobility	72	30.3%*	14	15.7%*
Self-Care - Community Management	24	10.1%	10	11.2%
Productivity - Paid/Unpaid Work	11	4.6%	5	5.6%
Productivity - Household Management	41	17.2%*	13	14.6%*
Productivity - Play/School	0	0.0%	2	2.2%
Leisure - Quiet Recreation	9	3.8%	13	14.6%*
Leisure - Active Recreation	28	11.8%	6	6.7%
Leisure - Socialization	8	3.4%	8	9.0%
Total	238		89	

Note. * represents those activity categories which both groups selected more frequently as being of concern or difficult to perform.

Life Roles of Clients in the Diagnostic Groups

Clients' perceived future life roles are accurate reflections of the roles they will play and also the activities associated with those life roles after their discharge from the hospital. For example, the role as a "home maintainer" would normally require a client to perform self-care and household tasks.

In this study, the life roles of clients in the orthopedic and stroke groups were analyzed by requesting the clients to report on their future roles using the Role Checklist. The results revealed that there were no significant differences ($p > 0.05$) between the profiles of life roles of clients in the orthopedic and stroke groups (Table 5.8). The findings also suggested that, with two exceptions, the male and female clients had similar profiles of life roles. A significant higher proportion of females (28.6%) than males (0.0%) perceived themselves as "volunteer" (*Chi-Square*=3.95, $df=1$, $p < .05$). Likewise, a significantly higher proportion of females (57.1%) than males (18.2%) indicated they would participate in religious activity - "religious participant" (*Chi Square*=4.82, $df=1$, $p < .03$).

In general, the life roles that were commonly perceived as potential future roles were, in descending order, "friend", "family member", "home maintainer", "hobbyist/amateur", and "care giver". The majority of the clients identified one or two out of these five future roles. Five clients predicted they would likely play all five roles at one time. The life roles that the clients were less likely to perceive playing in future were as "student", "volunteer", "worker", "participant in organizations", and "religious participant". The role patterns revealed in this study are consistent with other research on role changes among people with disabilities, particularly the roles of "home maintainer" and "family member" (Hallett et al., 1994). These activities are indoor activities rather than those which involve substantial traveling. The activities related to paid/unpaid work and play/school were not commonly considered by clients as future life roles. This coincides with their less perceived roles as volunteers, workers, and participants in organizations.

The clients' future life role profiles suggest that the clients would likely engage in activities which are related to self-care, such as personal care (role of "family member") and functional mobility; productivity, such as household management (roles of "home maintainer" and "care giver"); and leisure, such as quiet recreation ("hobbyist/amateur") and socialization activities ("friend and family member") as determined using the COPM. Comparing the results from the Role Checklist with those activities obtained from the COPM provided evidence regarding the substantive validity of the COPM. The notion that activities of concern or performed by a person should commonly be reflected in their life roles (see Chapter II) suggests a convergence between the responses made by the clients in the COPM Problem Definition and their perceived future life roles on the Role Checklist.

Table 5.8

Comparison of the Proportions of Clients Perceived Future Life Roles between Different Gender and Diagnostic Groups

	Gender		Diagnostic Groups	
Life Roles of Clients	Male (%)	Female (%)	Orthopedic (%)	Stroke (%)
Friend	100.0	92.9	96.7	88.9
Family Member	81.8	85.7	80.0	100.0
Home Maintainer	63.6	89.3	83.3	77.8
Hobbyist/Amateur	63.6	82.1	76.7	77.8
Care Giver	45.5	53.6	50.5	55.6
Religious Participant	18.2*	57.1*	46.7	44.4
Worker	45.5	17.9	23.3	33.3
Participant in Organizations	18.2	35.7	30.3	33.3
Volunteer	0.0*	28.6*	16.7	33.3
Student	9.1	21.4	13.3	33.3

Note. * represents the life roles in which significant differences ($p < .05$) were found between the proportions of male and female clients selecting them.

For each of the ten roles considered, a Chi-Square test was used to compare the activity patterns between clients who perceived themselves playing a particular life role and those who did not (Note: Here and elsewhere, the Chi-Square values were computed from non-independent frequencies, i.e. each client generated many responses. Thus all Chi-Square values should be treated as approximate indicators.). Each activity pattern was composed of the nine activity categories specified in the COPM (refer to Table 5.7 for details). The Chi-Square analysis hence was a 2 x 9 design (Yes/No - by - Activity) with $df=8$. No significant relationship ($p>0.05$) were found between the activity patterns of the clients and their perceived role for each of the ten life roles. Further inspection of the frequency distributions in the 2 x 9 contingency table of each life role revealed that the many of the frequencies in the cells were less than 5. These jeopardized the assumptions of using the Chi-Square test. Therefore, additional statistical analyses were conducted by clustering the nine activity categories into three areas, self-care, productivity, and leisure. A 2 x 9 design was then changed to a 2 x 3 design (Yes/No - by - Areas of Occupational Performance) with $df=2$. Results of the analyses are presented in the following sub-sections with respect to the different future life roles of clients.

Perceived Role of Student

There were seven clients (17.9% of the total sample) who perceived their future roles as students. Table 5.9 shows the patterns of activity of clients between those who perceived themselves playing a student role ("Yes") and those who did not ("No"). The results revealed that the seven clients who perceived a future student role less frequently (40.5% compared with 60.7%) were more likely to identify self-care tasks as more difficult to perform than those who did not (*Chi-Square*=9.67, *df*=2, $p<.01$).

Table 5.9

Patterns of Activity and Role of Student

	Perceived Role of Student	
Areas of Occupational Performance	Yes (%) (*n*=7)	No (%) (*n*=32)
Self-Care	40.5	60.7
Productivity	28.4	20.2
Leisure	31.1	19.0
Number of Activities Identified	74	252

This suggests that the clients who saw a future role as a student were less likely to perceive problems in performing self-care tasks when they were discharged from the hospital. The seven clients also put more emphasis on their ability to engage in leisure activities that did the clients who did not perceive a student role (31.1% versus 19.0%). Clients' activity patterns were consistent with their perceived roles as students who tend to have a more active life style than those who do not perceive such a role .

Perceived Role of Worker

Similar to the patterns revealed in the role of a student (Table 5.10), clients who perceived a future role as workers ($n=10$) showed less concern (45.5% versus 60.8%) with self-care activities than those who did not see themselves as workers (*Chi-Square*=7.63, *df*=2, $p<.02$). The same group of clients were more concerned (30.3% versus 18.5%) with their performance in the productivity area such as household management (e.g. "doing laundry"). Clients who perceived themselves as workers may expect that independent functioning is a pre-requisite to taking on the role as a worker. As a result of this perception, they expressed fewer concerns in self-care when compared with those who did not play such roles.

Even when perceiving a future role as a worker, clients still showed more concern about possible problems performing self-care activities than performing the productivity area. This is a common concern in an elderly population which reflects the ages of the study group (mean age 64.5).

Table 5.10

Patterns of Activity and Role of Worker

	Perceived Role of Worker	
Areas of Occupational Performance	Yes (%) ($n=10$)	No (%) ($n=29$)
Self-Care	45.5	60.8
Productivity	30.3	18.5
Leisure	24.2	20.7
Number of Activities Identified	99	227

Perceived Role of Care Giver

There were 20 clients (51.3%) who perceived themselves as future care givers (Table 5.11). Similar to those who perceived themselves as workers, clients identified significantly fewer concerns with their performance in self-care activities after their discharge (50.3% versus 64.9%) than those who did not anticipate a role as care givers (*Chi-Square*=7.20, $df=2$, $p<.03$). Again, clients may have been more likely to anticipate that independent performance in self-care is a prerequisite to a role as a care giver.

Table 5.11

Patterns of Activity and Role of Care Giver

	Perceived Role of Care Giver	
Areas of Occupational Performance	Yes (%) (n=20)	No (%) (n=19)
Self-Care	50.3	64.9
Productivity	24.1	19.1
Leisure	25.6	16.0
Number of Activities Identified	195	131

Perceived Role of Home Maintainer

There were 32 clients (82.1% of the total sample) who perceived a future role as home maintainers (Table 5.12). Significant differences were observed in the patterns of activity in the three areas of occupational performance between those who did perceive themselves as home maintainers and those who did not perceive themselves as such (*Chi-Square*=5.84, df=2, $p < .05$). Similar to the roles of worker and care giver, the 32 clients who perceived themselves as home maintainers identified fewer concerns with their performance in self-care than those who did not. In contrast, more concern was shown for performance of productivity activities (24.3% versus 8.7%) among the clients who saw themselves in the role of home maintainers.

Table 5.12

Patterns of Activity and Role of Home Maintainer

	Perceived Role of Home Maintainer	
Areas of Occupational Performance	Yes (%) (*n*=32)	No (%) (*n*=7)
Self-Care	55.0	63.0
Productivity	24.3	8.7
Leisure	20.7	28.3
Number of Activities Identified	280	46

Perceived Role of Hobbyist / Amateur

There were 30 clients (76.9% of the total sample) who perceived themselves as hobbyist/amateur (Table 5.13). Significant differences were found in the patterns of activity of concern to the clients (*Chi-Square*=6.23, *df*=2, $p < .04$). Similar to the patterns revealed in other life roles, the 30 clients identified fewer concerns with their performance in self-care activities and more concerns with productivity activities. Furthermore, clients who perceived a future role as a hobbyist/amateur showed more concerns with their performance in leisure activities (such as quiet recreation) than those who did not perceive themselves as a hobbyist/amateur (23.9% versus 7.1%).

Table 5.13

Patterns of Activity and Role of Hobbyist/Amateur

	Perceived Role of Hobbyist /Amateur	
Areas of Occupational Performance	Yes (%) ($n=30$)	No (%) ($n=9$)
Self-Care	54.9	64.3
Productivity	21.1	28.6
Leisure	23.9	7.1
Number of Activities Identified	284	42

Other Life Roles

No significant differences were revealed in the patterns of activity of the clients who perceived their role as friends (94.9% of total sample), family members (84.6%), religious participants (46.2%), participants in organizations (30.8%), and volunteers (20.5%). For the roles of friends and family members, these insignificant differences may be attributable to the large number of clients who perceived these particular roles. For instance, 37 out of 39 clients chose the role of "friends". The activities that they selected on the COPM occupied 96.9% of the total clients' responses. Consequently, the differences in the proportions of the different areas of occupational performance were difficult to see.

Insignificant differences in patterns of activities between clients who did and did not perceive their roles as religious participants, participants in organizations and volunteers may be explained by the confusing operational definitions of those roles (as specified in the Role Checklist). For example, "participant in organizations" was defined as, "Involvement, at least once a week, in organizations such as the American Legion, National Organization for Women, Parents without Partners, Weight Watchers, and so forth" (Oakley et al., 1986). The term "involvement" inadequately reflects the pre-requisite demands in order for the clients to claim the role of "participant". For instance, "involvement" can mean being physically present at every event held by the organization, or it can mean being interested in an issue but without any physical involvement. Consequently, it is difficult to define sets of activities that might be associated with the role of a participant in an organization, such as functional mobility and community management. Because of this, clients may not be able to make consistent judgments on whether they should perceive themselves as becoming involved in the organization or not, and hence the patterns of activity required for performing such a role are not clear.

Summary of Pattern of Activity and Perceived Future Life Roles

The notion that clients' life roles and patterns of activity are related constructs has provided the basis of establishing convergent validity between the activities selected by the clients on the COPM and the clients' future life roles as revealed from the Role Checklist. Results suggest that clients' responses on the COPM reflect their concerns with different patterns of activity. These concerns are consistent with the particular life roles that the clients perceive themselves playing in the future. Except for several life roles identified as exceptional in this study, the majority of the life roles are associated with the relevant patterns of activity. This has provided evidence on supporting the substantive component of construct validity of the COPM. When the clients were asked to select the activities that they perceived as being difficult to perform after their discharge, the activities did not appear not to be randomly selected. Rather, it seemed that a more complicated process was undertaken through which clients considered their life roles, and the responsibilities and activities associated with these roles before decisions were made on selecting particular activities on the COPM.

Analyses of COPM Testing Processes

The administration of the COPM, like other clinical instruments, assumes that clients performing the same tasks go through the same processes and hence provide consistent responses. However, it is argued that although the testing procedures of the COPM are largely standardized, clients who were assessed with the instrument were likely to interpret tasks and processes differently. As a consequence, clients' responses were likely to differ as well as therapist's clinical interpretation and inferences.

The first portion of analysis in this section focused on the characteristics of clients' responses in the Problem Definition in the three areas of occupational performance in the COPM. In the second portion, evidence is presented on the reasoning behind the responses made in rating the Importance of the activities identified, rating the self-perceived Performance in those activities, and then rating the Satisfaction with performance. The data analyzed in this part were drawn from the semi-structured interview with 21 clients in the retest occasion either within 3-days (orthopedic group) or 7-days (stroke group) after the date of the initial assessment. After the clients completed the COPM, they were requested to describe retrospectively the processes that they experienced and remembered as they went through the tasks involved in completing the test. This included clients' thoughts, feelings, impressions, and methods of self evaluation. The protocol analysis is believed to reflect the representation of the construct of occupational performance in both the structure and processes of using the COPM as a clinical assessment tool.

Problem Definition of Clients

After the retest of the COPM, the clients were asked to tell how they interpreted the term "problem". They were also asked to describe how they had reached the decisions on identifying those activities as problematic. The focus here was on the processes through which responses were derived rather than on the content of those responses (as discussed earlier).

When the clients were interviewed by clinicians in the retest, the following leading questions were asked,

> "When I asked you to tell me the problems that you think you might have after you are discharged from the hospital, what were the issues that came immediately to your mind?"
>
> "Was there anything particular that you thought of when you said (a particular functional problem raised by the client) was a problem for you?"

In general, the clients were cooperative during the interview. Most of them were willing to reveal their feelings and experiences to the clinicians.

Meanings of "Problem"

According to the Test Manual (Law et al., 1991), the step of Problem Definition requires the clients to be interviewed "to determine if they are having any problems in occupational performance" (p.13). In this study, clinicians prompted clients by asking: do you need to, do you want to, or are you expected to do...? To proceed further, clinicians asked: can you do, do you do, or are you satisfied with the way you do...? These dialogues were thought to encourage clients to identify their problems in occupational performance in the context of role expectations, perceived performance of those activities, and satisfaction with the performance.

When clients were asked to identify their problems, data indicated that most of them went through an objective and systematic problem identification process. The common strategy that clients seemed to use was to compare their present abilities with their abilities before the hospitalization. The following are some of the responses made by these clients,

> "I based this on my previous experience, and I compared my abilities with my previous performance."
>
> I thought about the way that I do it right now. I tried to compare the present performance with the past."

Responses from four clients indicated that the comparisons made between the past and present performance were quite systematic and thorough. For example,

> "I considered the stairs first because I think of the architecture of my house. There are stairs which I can not get around at home."

Another client said,

> "I mentioned the bed but the next thing I thought of was getting into the house. Why worry about the bed if I can not get into the house?"

Besides this rather systematic method of problem identification, some of the clients interviewed used a strategy which relied more on their emotion and intuition:

> "I want to go home and take it easy. I also want to play sports. However, I have a weakness in right arm and leg....I define something as having a problem when I am not able to take care of myself and something that is difficult for me to overcome."
>
> "I am being confined to a house....I can not think of any problems right now. There may be other things that come up."

Obviously, the two strategies of identifying problems were very different. The responses of the clients showed that those clients who more systematically analyzed their situations came up with more specific functional problems than those using a more general and intuitive strategy. Some of the examples of the activities selected by clients using the former strategy were: cooking and preparing meals, putting on/taking off shoes, going outdoors for coffee, and cutting toe nails. These were very specific both to their roles and household duties. In contrast, the activities selected by the clients with a more intuitive strategy were: walking, outdoor activities, and work. These were broad and less role-specific. As a result, different strategies seemed likely to produce different patterns of activity, and probably a variation in the numbers of activities identified. Therefore the assumption that similar responses are elicited in the problem identification process of the COPM is challenged.

Another criterion adopted by the clients to define "problems" was also noteworthy. Generally speaking, clients did not perceive an activity as a problem even though they could not perform it if there was somebody who or something that could offer assistance. Such people included spouse, friends, sons or daughters, and other relatives. One client responded,

> "I will get support from my family members and friends....I have friends to cook for me, and my family will help with shopping and vacuuming....I could get the method to prepare everything and adapt to the environment."

Another client mentioned,

> "I would not be able to manage by myself but because my grand-daughter will be with me, I will be alright."

In fact, the support system that the clients had was a predominant factor in determining whether a "problem" was recognized as such in the COPM. This finding reveals that clients considered both their occupational performance and sociocultural component when they made selections in the Problem Identification section in the COPM. As one client commented,

> "If I get the tools (equipment) then it would be much easier for me to get around in the house and be independent. I need to consider other people's situation (client's children). They all have small children. I have been living alone for 24 years and it is nothing new. I will get used to it and not feeling [sic] bad about it."

This quotation reflects the fact that when clients identified problems in the COPM, the test taking processes were far more complicated than anticipated. Decisions were made following objective or subjective and specific or general analyses of the problems. The links between occupational performance and specific performance components, especially sociocultural, were very clear. The processes by which clients identified the problems for subsequent rating in the COPM were diverse and complex.

The findings showed that clients seemed to understand the instructions given in the COPM. The fact that clients considered both occupational performance and performance components at the same time in making responses to the COPM is consistent with the speculations in the model of occupational performance. However, whether these complex processes were based on the natural thinking processes of the clients or on the prompting by the clinicians requires further investigation.

Rating of "Importance" of Activities

The discussion in this section focuses on the ways that clients interpreted the concept of "importance". According to The Manual (Law et al., 1991), the clients were asked "to rate the activity in terms of its importance in his or her life" (p.15). In obtaining the protocols in this study, clients were prompted to reveal the processes through which they assigned ratings using the Importance scale and the rationale which they used to evaluate the importance of these activities (identified in the Problem Definition) in their life. The following are examples of the leading questions used by the clinicians during the retest interview,

> "When I showed you this card (Importance), what did the word "importance" mean to you? When you called something "important" to your life, which aspects did you refer to?"
>
> "For the activity, (select the activity with the highest rating), you gave a very high rating on the scale. How would the performance of this activity affect your life (or what would it mean to you)?"

Similar to the "Problem Identification" section, questions under the "Importance" category were asked after the clients completed the entire COPM.

Meanings of "Importance"

The analysis of clients' protocols indicated that clients showed understanding of the concept of "importance" and the use of the Importance scale. The clients' words and phrases revealed that the clients perceived the term "importance" at two levels, one abstract and emotional and the other concrete and objective. Descriptors at the emotional level included: "high priority in life", "vital to my well being", "anything tied with my independence", and "enjoying myself". In contrast, descriptors at the objective level included: "basic", "necessary", essential", "no other choice" and "I know I need it, I must have it". Protocols at both levels revealed similar understanding of the term "importance":

> "Those are the basic things that I do everyday, such as dressing, toiletting, and showering. They are essential." (objective level)

"Going out means a lot to me; it means I can do anything if I could go out." (emotional level)

"Dressing. I have to be able to dress myself; it makes me independent. I want to be just like I was before I came in here." (emotional level)

The different descriptors used likely reflect the differences in clients' educational background, value systems, life experiences, abilities to express themselves, and nature and types of the activities.

The descriptors related to the term "unimportance" were also extracted from the clients' protocols. They, too, were divided into an emotional level and an objective level. Some of the emotional descriptors used included: "it will not change my life" and "it would not be the end of the world". Examples of objective descriptors include: "I could do without it", "I would hire somebody to do it", "if the problem is temporary", and "I would use other methods to do it". Most of the clients suggested that activities performed with a low "urgency" or "frequency" were usually regarded as "unimportant".

Findings from the protocol analysis indicated that clients had an adequate understanding of the concept of "importance". This was reflected in the ways clients interpreted the two ends of the "importance" spectrum: "important" and "unimportant".

Discrimination of Importance Scale

Although the protocol findings suggested the clients' general understanding on the concept of "importance", this did not imply the meaningfulness of COPM's Importance scale to the "importance in life" construct. The 10-point Importance rating scale ranged from "1" - "not important at all" to "10" - "extremely important". The analysis of the clients' protocols were used to discern the meaningfulness the clients attached to the different numbers on the Importance scale; i.e., to see if the different numbers did in fact reflect different levels of "importance" to the clients.

Expressions of the clients like "very important", "necessary", "highest importance", and "enjoying doing it" were associated with the ratings of "9" and "10". In contrast, lower ratings such as "5" and below were associated with "not a priority to me", "don't expect to do it for a while", and "a temporary problem". For example, one client explained,

> "Being clean is very important to me (rating of 9), so laundry is one of the activities which is very important to me. Going to the university and taking courses is important because I want stimulation of my mind (rating of 8). However, I could live with my sister and brother who could help me out (rating of 5)."

Another client reported,

> "I gave a 10 to yard work because I enjoy it so much. I gave a 4 to cooking which is a low rating because it is not a priority to me. I don't think about it till I get hungry. I would rather be out in the garden."

Evidence showed that the discrimination of "important" and "unimportant" were well anchored along the Importance scale. However, there were inconsistencies within smaller increments such as ratings of 10 and 8.

Protocols seemed to suggest considerable confusion among some clients using the Importance scale especially in the middle range ratings. A rating of 8 was found to be problematic as illustrated below,

> "I gave bathing a 10 which means 'darn well' when I have a bath. I gave cleaning (household) an 8 because my friends help me clean my house."

Another client expressed similarly,

> "I gave leisure an 8 because I can eliminate it without great impact on my life. Work got a 10 because I need to go back to work to support my family."

The findings seem to suggest that the scale was effective in locating those clients who rated an activity as "important" or "unimportant" in life, but not the "less important" ratings such as between 6 to 8. Except for the ratings of 10 or 9 (important) and 5 or 4 (unimportant), other points the clients exhibited on the 10-point Importance scale seemed difficult to differentiate. Consequently, the discrimination power of the 10 points along the ordinal scale was challenged.

Ratings of Perceived Performance

The clients rated their performance of each of the five most important problems they had using the Performance rating. According to the Test Manual (Law et al., 1991, p.16), the 10-point Performance rating scale was used to obtain a "subjective evaluation" of clients' "current performance" in the selected activities. The following analysis is structured to investigate the meaningfulness and usefulness of the Performance scale in reflecting the self-reported functional performance of clients. In obtaining protocols, clients were prompted to reveal the processes through which they assigned ratings using the Performance scale and the rationale with which they evaluated their own performance in the activities they had selected. Examples of the questions used by the clinicians to elicit responses are as follows:

> "When I asked you (the client) to rate your own performance, what were the things that you considered before you gave me the answers?"
>
> "How did you come up with (a rating of performance) on (a particular activity)?"

Other prompts, such as what a high (9 or 10) or low (4 or 5) performance rating meant and what particular factors affected their performance of the activities were also used.

Meanings of "Performance"

The analyses of the protocols revealed that most of the clients interviewed were consistent in their interpretation of "Performance". Moreover, the clients' descriptions suggested that their ratings against the Performance rating scale were meaningful. These findings appear to converge with the conclusions drawn by other studies on validity of self-report information such as Elam et al. (1991) and Myers, Holliday, Harvey, and Hutchinson (1993).

The analysis of the protocols revealed that clients generally interpreted the term "Performance" in two ways. One group of clients viewed "Performance" as a dichotomous concept. They regarded themselves as either in a "can do" or "cannot do" category. Another group of clients added a qualitative component to their interpretation. They included "how well" they could perform the activity. Below are some of the descriptions from the clients in the "can do" category,

> "When I say performance, I mean I am capable of doing them myself without help or support."
>
> "Performance means something that is done...It also means my ability to do something."

The second group of clients added descriptors such as "do it well", "difficulty level of tasks", "comfort", "speed" and "efforts" to qualify their performance. For example,

> "Performance means whether I am able to do things, and how well I did it."
>
> "I hope that I could do stuff normally at an average rate and do what others are doing. I hope that the slowness is temporary."

The responses indicated that clients had a clear concept of "performance" which was similar to the concept as defined in the Test Manual.

Discrimination of Performance Scale

The Performance scale defines its two extremes as "able to do it well" and "not able to do it". However, during the interview with the clients, it was found that the dichotomous concept of "performance" revealed by some clients was likely attributable to their modesty in expressing themselves. No clients rated their performance as 10 on the Performance scale in the initial assessment. Instead, most of the ratings were within the 1 to 5 range.

Findings suggest that most of the clients had a clear concept of self-rated "Performance". Clients used the Performance scale to rate their performance in daily activities in a very discriminating fashion. These ratings were determined by the availability of "adapted methods", "adaptive equipment", "assistance", and "efforts required" when the activities were performed. In fact, these qualifying expressions used in the protocols illustrate the meaningfulness of the 10-point rating scale. For example,

> "If I have the appliance (adaptive equipment), to go to bathroom, the rating is an 8. However, if there is no appliance, I would give a 1, because I could not do it in the first place....A 5 to me is at the borderline of doing something whether it is with comfort or a waste of energy."
>
> "I gave cooking a 5 because I think that, at present, I am okay with cooking. For yard work, it was a 1 because I can not do it at all."
>
> "It is 1 because right now I could not do anything. I get a 7 since I can print and it is legible but it takes longer."

The protocols clearly suggest a systematic evaluation going on when clients were asked to rate their own performance. Further evidence shows that the scope of their considerations went beyond their hospital experiences. This demonstrates that some of them had done a very thorough analysis of their home and the hospital environment and hospital during their hospital stay/confinement and after their discharge. As one client revealed,

> "I think I can do it (committee work) by remote. I can do it from home and I give myself a 5."

Another client said,

> "In and out of bed, I gave a 7 because I've got a little ways to go yet. I still have a little trouble. I am not 100%. For stairs, I gave it a 1 because I cannot put weight on my left foot. It worries me how I will learn to do stairs."

When compared with the Importance scale, the Performance scale was more meaningful and refined because evidence from clients' protocols indicated that they had good understanding of the concept of "performance". Ratings of performance against the Performance scale were found to reflect clients' functional level of particular activities. Self-rating of clients performance on the COPM were shown to be appropriate and meaningful.

Rating of Satisfaction with Perceived Performance

After rating the perceived performance, clients were requested to rate their satisfaction with the perceived performance using the 10-point satisfaction rating scale. The Manual (Law et al., 1991) stated that this scale was used to measure the "subjective evaluation" of clients' satisfaction with their current performance. In obtaining the protocols, the clients were prompted to describe the processes through which they analyzed their own performance and assigned their satisfaction ratings. Their interpretations of the term "satisfaction" were also explored. Some examples of the questions asked in this part are,

> "Look at the Satisfaction card which I showed you before. What does the word "satisfaction" mean to you?"
>
> "What considerations did you have when you rated your performance in (an selected activity) as (respective satisfaction rating)?"

Meanings of Satisfaction

The clients' protocols revealed the vagueness with which "satisfaction with performance" was interpreted. "Satisfaction" seemed to be a more difficult construct to understand and quantify than "Performance" and "Importance". The analysis of protocols indicated that there were several ways in which clients described their "satisfaction with performance". One interpretation of the term "satisfaction" seemed to be the results of comparisons between "what one has now" with "what one had before". Phrases related to "happiness", "happy with one's ability" and "feeling content" were used:

> "For walking, I just feel I like a two-year old starting to learn how to walk again. I can not do much better than this. I am not very happy now. For cooking, I can not open cans. I mean there are a lot of things that I can not do now. That is why I am not satisfied with it. I ask for an electric can opener."

Another interpretation of "satisfaction" was largely related to the notion of "further improvement" and "potential of getting better". The following descriptions reflected this concept:

> "I find myself progressing everyday. I know that it will heal up one day. So, I am quite satisfied with whatever function that I have got although I am still pretty disabled."

> "I define satisfaction as 'content'; it means something that makes you happy....What do you expect an old man like me to do in the hospital and getting good progress."

Although it was implied, most of the clients' responses suggested that "expectation" and "hope" were important factors influencing their evaluation of how satisfied they were with their performance. The notion of personal "expectation" was clearly shown by one client:

> "I interpret the term "satisfaction" as am I doing it to the level I think I should be doing it. I will be satisfied when I can do the things that I used to do."

Quite a few clients also mentioned that satisfaction was highly related to or even equivalent to the ability to perform activities:

> "Satisfaction means being able to walk. I can not go down the street and to the neighborhood. I do not want to stay and be put in one place. I haven't been outdoors yet...I can not say I am satisfied except if I could do all of them well."
>
> "I tend to link satisfaction with my ability to do things. It all depends on what I can do."

Findings from the analysis suggested that the concept of "Satisfaction with Performance" was complex and individualized. Clients revealed different perspectives in judging their own satisfaction with their perceived performance.

Protocol analysis has shown that "satisfaction" is a more complicated construct than "Importance" and "Perceived Performance". However, this is not a surprising result as it matches very well with theories of well-being and life satisfaction in which the construct of "satisfaction" is essentially multi-dimensional (Chamberlain, 1988; Meadow, Mentzer, Rahtz, & Sirgy, 1992). The judgment theories proposed by Carp and Carp (1982) and Michalos (1980, 1985) suggested that satisfaction of persons at any given time is a function of a cognitive comparison between the actual performance and some standards or conditions. In the conceptualization of "satisfaction with performance" in this study, clients compared their perceived current performance with: 1) their expected performance (standard); 2) their previous performance (standard); and 3) actual conditions or performance (condition). The appropriateness of using the Satisfaction scale to measure clients' satisfaction with current performance in the COPM hence is evident.

Discrimination of Satisfaction Scale

Although clients' conceptualization of "satisfaction" largely converged with that of the COPM's scale, the protocol analysis revealed some problems in the meaningfulness and usefulness of the "Satisfaction" ratings especially within the middle range of the scale. As one client explained his different "Satisfaction" ratings,

> "Satisfaction means I am quite happy with my ability. Yard work is 1 because I cannot do anything but look and that bothers me though I know it will be a temporary thing. Cooking I gave a 5 because I like to cook."

Another client revealed her "Satisfaction" ratings as,

> "Satisfaction is the degree of accomplishment. I gave a 6 to in/out of bed. I noticed a small improvement lately so my satisfaction is getting up. I gave 1 means [sic] it is impossible for me to even start."

The above two examples show that clients used different criteria in making decisions on quantifying their degree of satisfaction on the scale. The former rated activities in terms of both ability and enjoyment, whereas the rating of the latter client was based on her progress in performing an activity.

Problems with meaningfulness of the "Satisfaction" ratings were not restricted to those within the middle range. Clients were found to interpret the rating of 1 (not satisfied at all) differently,

> "I say I am satisfied when I am able to do something. When I assigned 1, it means that it is not easy to get in the tub with my surgery and my bad knee."

Another client indicated,

> "Satisfaction means how happy I am with how I do...For shower, I assigned a 1 because it is so difficult and it is something that I really want to do. I have to 'psyche' myself up all the time."

Besides the different interpretations of “satisfaction” ratings, another problem was found when clients used their current performance as a referent for rating satisfaction instead of comparing the performance with past or expected performance. In this case, their “Performance” and “Satisfaction” ratings were indistinguishable. As a result, the meaningfulness and usefulness of the Satisfaction scale in the COPM were very much challenged.

Conclusion on Substantive Validity of the COPM

In this section, substantive related evidence of the test content and testing processes of the COPM was presented and discussed. To begin, the truthfulness of the activities selected by clients was determined by establishing the convergence of clients’ patterns of activity and their perceived life roles. The type of activities identified by the clients on the COPM were found to be relevant to the demands that were likely to be incurred from clients’ perceived life roles. Analysis of clients' protocols suggested that systematic and complex processes were used by the clients to identify the activities which they perceived to be problematic. It was evident that clients considered performance components and occupational performance dysfunction before responses were made.

Protocol analyses also revealed that the clients interpreted differently the constructs of "Importance", "Performance", and "Satisfaction". The processes that the clients went through before ratings were assigned using the respective three subscales in the COPM reflected good evidence of substantive validity of the instrument. However, the evidence supporting the meaningfulness of different numerical ratings showed weaknesses in each subscale. As a consequence, the appropriateness of using the 10-point rating scales to measure the three constructs in the COPM was challenged, especially for the Satisfaction scale. In the next section of this chapter, more evidence will be presented that relates to these challenges by exploring the structural fidelity of the "Performance" and "Satisfaction" subscales.

Section C - Structural-Related Evidence

In the previous two sections of this chapter, the content-related and substantive-related evidence of the COPM were presented and analyzed. To accumulate further evidence of construct validity and strengthen the inferences that can be made about clients' occupational performance from the scores on the COPM, this section focuses on the dimensionality and scoring system of the COPM. Before this structural-related evidence is presented, the results of estimating the stability of test responses of the COPM are reported. The stability estimates were intended to contribute to the knowledge about psychometric properties of the COPM. They also shed light on the strengths of relationship that the COPM scores would yield with other instruments in establishing criterion-related evidence in the present study.

The structural-related evidence is divided into two parts each reflecting the different method used in establishing the evidence. The first part is exclusively quantitative, as it explores the structural fidelity of the COPM in terms of inter-item relationships and dimensionality. The second part is a qualitative review of the ratings and comments by members of the panel review on the structure of the COPM.

Short Term Stability of COPM

From the 39 clients included in this study, 21 were selected for the retest (instead of 30 clients as planned), with 15 from the orthopedic group and six from the stroke group. The average number of days from the initial assessment to the retest was 2.7 (SD=0.7) days for the orthopedic clients and 4.0 (SD=1.4) days for the stroke clients.

In the retest, only the COPM was administered. Different from the original plan of the study, the therapists who implemented the reassessment to the clients were different from the therapists who conducted the COPM in the initial and pre-discharge assessments. This arrangement was a compromise solution to the problem of therapists being too busy in their clinical work to carry out the assessments on all three occasions.

Notifications for the retest of clients were sent by the case therapists after initial assessment to a team of two therapists responsible for the re-assessment. After receiving the notifications, clients were re-assessed and interviewed unless they were medically unfit to participate. The COPM was administered to the clients by the researcher and one other therapist employed by the hospital.

This design made the estimation of the short term stability contaminated by the errors associated with differences among the testers. Moreover, the changes within the clients between the initial assessment and the reassessment due to the treatment effects and natural recovery contributed additional errors to the stability of the test scores between the two assessment occasions. It was impossible to isolate which portions of the errors were attributable to which of these two sources. These problem meant that the concept of test-retest reliability as proposed in the present study was inappropriate. Consequently, results of this study were at most indicators of stability of clients' responses over a short period of time.

Results of clients on the COPM in both the initial assessment and reassessment are summarized in Table 5.14. Differences were found in the mean Performance and Satisfaction subscale scores between the two assessment occasions. However, the differences were not statistically significant (Performance subscale score: $t=-1.46$, $df=20$, $p<.16$; Satisfaction subscale score: $t=1.04$, $df=20$, $p<.31$). The same insignificance was revealed after the clients were grouped with respect to their diagnoses. These insignificant findings can be explained in part by the small sample size. Another explanation of these findings is the large standard deviations of the COPM subscale scores which reflected large variability among individual clients in the same diagnostic group. The short term stability of the COPM subscales scores using Pearson's r are the correlations presented in Table 5.14 which were .32 for the Performance subscale and .09 for the Satisfaction with performance scale.

Table 5.14

Comparison of the COPM Mean Performance and Satisfaction Scores between Initial Assessment and Reassessment

	Performance Subscale		Satisfaction Subscale	
Diagnostic Group	M	SD	M	SD
All (*n*=21)	23.1 (28.0)	12.1 (14.4)	24.2 (20.3)	14.6 (10.9)
Orthopedic (*n*=15)	20.7 (24.1)	11.5 (13.0)	22.9 (16.5)	15.4 (9.9)
Stroke (*n*=6)	29.0 (38.0)	12.4 (13.5)	27.7 (29.6)	13.2 (7.8)
Pearson's *r* between Assessment Occasions	.32		.09	

Note: Means and standard deviations inside parentheses were obtained in the reassessment occasion.

The scatterplots of clients' initial versus retest COPM subscale scores are presented in Figures 5.1 (Performance subscale) and 5.2 (Satisfaction subscale). The distributions shown in the two scatterplots confirmed low correlations of clients' COPM subscale scores between the initial assessment and retest occasions.

Figure 5.1

Scatterplot of the COPM Initial and Retest Performance Subscale Scores

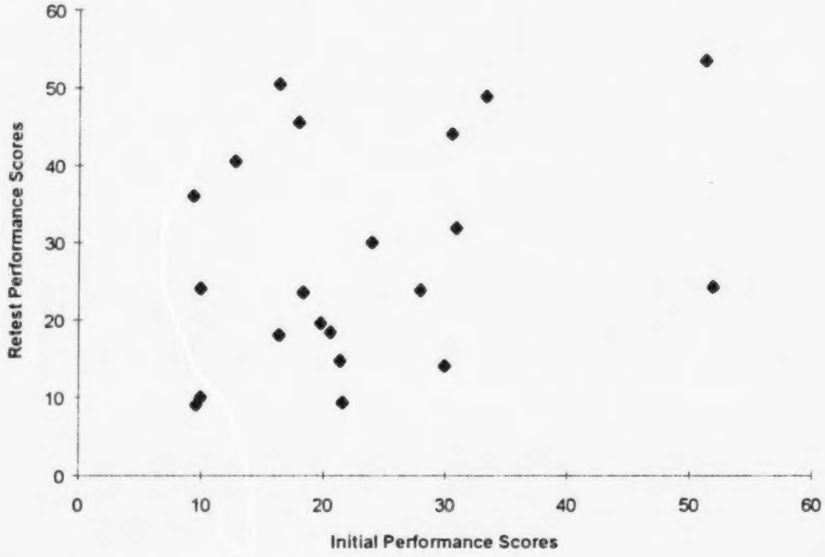
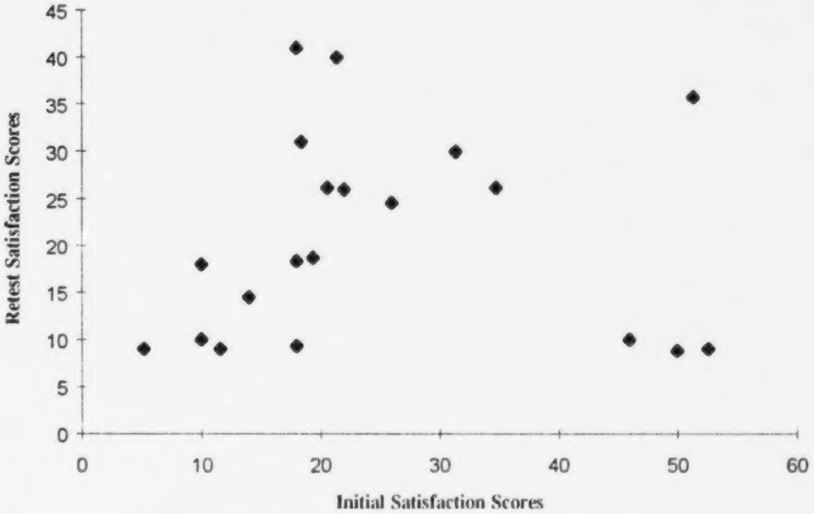


Figure 5.2

Scatterplot of the COPM Initial and Retest Satisfaction Subscale Scores



The low correlations for both subscales of the COPM indicated low stability of the COPM scores between two assessment occasions. As mentioned earlier in this section, the differences in therapists administering the COPM and the changes within the clients between the initial assessment and reassessment occasions accounted for the two major sources of random errors in the present stability estimates.

Tester and Occasion Effects on COPM Problem Identification

One of the major flaws in the short term stability in this study was the use of different therapists in conducting the COPM in the initial and retest occasions.

The second flaw was the changes in the clients between the period from the initial assessment to the reassessment occasions. In order to understand the extent to which clients' responses changed in the two assessment occasions, an analysis was conducted to explore differences in the pattern of activities selected by the clients in the COPM.

Table 5.15 shows the comparisons of the patterns of activities identified by clients in the two diagnostic groups. In general, clients in the reassessment occasion identified fewer activities than in the initial assessment occasion in the self-care, productivity, and leisure activities categories. The differences were found to be statistically significant in the orthopedic group but not in the stroke group. In the orthopedic group, the total number of activities identified by the clients as difficult to perform in the initial assessment was significantly higher than in the reassessment (t=2.07, df=13, p<.02). Similar results were revealed when the activities were further categorized into self-care (t=3.62, df=13, p<.01) and leisure (t=2.21, df=13, p<.04). However, the difference in the productivity category was not statistically significant (t=.92, df=13, p<.37).

Table 5.15

Pattern of Activities Selected by Clients between Initial Assessment and Reassessment

	Mean Number of Activities Selected by Clients			
	Orthopedic Group ($n=15$)		Stroke Group ($n=6$)	
Activity Categories	Initial Ass.	Retest	Initial Ass.	Retest
Self-Care	5.0 (1.9)	3.0 (1.9)	5.8 (2.5)	3.9 (1.3)
Productivity	1.9 (2.5)	1.2 (1.2)	2.7 (3.1)	1.5 (1.1)
Leisure	1.9 (2.1)	0.7 (0.6)	3.3 (2.2)	1.3 (1.5)
Total	8.9 (5.4)	4.9 (2.7)	11.8 (7.1)	6.7 (2.7)

Note: Values in parentheses are standard deviations of the mean number of activities.

In the stroke group, even though the numbers of activities selected by the clients at the reassessment were less than at the initial assessment, the differences were not significant. The lack of significance is largely attributed to the fact that there were only six clients who participated in the retest. The variability among individual clients on both occasions accounts for the failure to statistically demonstrate those differences.

Besides the number of activities, the nature of activities that clients identified as difficult to perform were found to be different between the initial assessment and the reassessment. The consistency of the activities identified in the two testing occasions was quantified by using the percentage of agreement index. The mean percentage of agreement between the initial assessment and the retest was 37.3% with a range from 0.0% to 62.5%. Results indicated that clients identified different activities in the two testing occasions. In general, when clients identified ten activities in the initial assessment, only three to four activities were consistently identified in the reassessment. When the identified activities were limited to five with the highest Importance ratings, the percentage of agreement dropped slightly to 36.2%, ranging from 0.0% to 60%.

Table 5.16 provides a sample of the activities identified by one client on the two assessment occasions. The percentage of agreement of the activities identified by this client between the initial assessment and the retest was 0.0%. Content analysis suggested that the client identified more problems in the areas of productivity and leisure on the retest than on the initial assessment. Despite the fact that the same number of activities were identified in the area of self-care (four activities), none of the activities identified in the initial assessment were re-identified in the retest. This constitutes a fundamental problem of the COPM when activities identified by clients as difficult to perform are inconsistent in the two different assessment occasions. This is thought to impact the short term stability of the COPM and hence the ratings of Importance, Performance and Satisfaction (even though the differences in the subscale scores between the initial assessment and retest were statistically insignificant).

Table 5.16

Activities Identified by One Client in Initial Assessment and Reassessment

	Activities Identified by One Client	
	Initial Assessment	Reassessment
Self-Care	Hold a knife to cut meat	Using a Toilet
	Wash Self with Towel	Bathing right side of body
	Walking	Getting in and out of bed
	Writing	Taking a public transit
Productivity		Grocery shopping
		Laundry
		Cooking a meal
		Cleaning house
Leisure		Watching movie

Results suggested that the clients made different responses on the COPM in the two assessment occasions. There was an evident reduction in number and type of activities selection. However, the design of this study did not allow further inferences to be made on whether the discrepancies were attributable to the variations among the clients, or the clinicians conducting the assessment, or the influence of treatments. Among the clients, assessments with the COPM on different testing occasions may have resulted from a difference in concerns, performance and satisfaction due to progress from the treatment program or natural recovery. In other words, clients' functional status may have changed during the period between the initial assessment and the reassessment. However, the activities of one client as shown in Table 5.16 did not support this speculation. In the area of self care, the difficulty levels of the activities identified by the client as difficult to perform between the two testing occasions were by and large similar. One could argue that the differences were attributable to the increase in clients' level of awareness as they developed more concern with their independence before and during the retest occasion. As reported previously, however, there was a general reduction of number of activities from initial assessment to the reassessment.

Table 5.17 shows the activities identified by another client. The number of activities was reduced from nine to five. Analysis of the activities also revealed that the activities identified on the two occasions did not differ in their difficulty levels. Hence the likelihood of change in the ability levels of the clients between the two test occasions was not strongly supported.

Table 5.17

Activities Identified by Another Client in Initial Assessment and Reassessment

	Activities Identified by Another Client	
	Initial Assessment	Reassessment
Self-Care	Washing back and lower limb	Bathing
	Dress lower limb	Dressing
	Getting in and out of bed	Toiletting
	Climb up and down stairs	
	Walking	
	Getting in and out of car	Driving
Productivity		
Leisure	Reading in sitting position	
	Going to opera	Going to theatre
	Visit friends	

To explain such phenomenon, the clinicians speculated that they may have put different emphases and used different approaches when conducting the assessment on the different occasions especially since the COPM is not "well standardized" in its testing procedures. During the semi-structured interview, clinicians were free to prompt and guide the clients. As a result, responses made by clients easily differed (as revealed in this section). However, observations of clinicians interacting with the clients on the COPM were not part of the validation study. Such a speculation hence is more a clinical inference from the researcher than a verification of the unreliability of the test. Future research on this area is strongly recommended.

Conclusion on the Short Term Stability of COPM

The inferences that can be made based on the stability estimates of the COPM were unfortunately limited by its small sample size and by the research design. Consequently, it is impossible to isolate the "occasion" from the "tester" effects in explaining the instability of the COPM scores. Sources of error could be attributed to the occasion effects, the tester effects, or a combination of both. The instability of the COPM subscale scores would lower the values of the inter-item correlations of the COPM to be presented later in this section and the values of the correlations between the scores of the COPM and other clinical instruments (criterion-related evidence) in Section IV of this chapter.

Structure of the COPM : A Quantitative Analysis

The COPM does not possess all of the structural characteristics of a typical clinical assessment instrument has. It does not have a specific pool of items which differentiate the test from the non-test domains. Each client has the possibility of selecting a set of activities in self-care, leisure, and/or productivity which would be totally different from that the set selected by another client. Evidence has shown that clients' activity patterns may differ according to their own life roles, needs, and immediate environment (see Section B, Chapter V).

In the subsequent analyses, only 34 out of the total 39 clients were included. Five clients were excluded because they identified fewer than five activities that were identified as difficult to perform.

COPM Activity Items

After the activities are selected, clients are requested to assign an Importance rating to each activity item. The Importance ratings are not changed throughout the assessment process including the pre-discharge occasion. The Importance ratings assigned to the selected activities become the weights used to obtain the item Performance and Satisfaction scores. In this section, "activity items" refer to the five most important activities nominated by the clients, that is, those given the five highest Importance ratings.

COPM Scoring System

The test domain of the COPM suggests a two-dimensional structure for the instrument. These two dimensions are "perceived performance" and "satisfaction with performance". Clients are asked to rate each activity item separately with two 10-point rating scales to reflect their perception of 1) their performance of the activity and 2) their satisfaction with their performance. Each score is separately computed to yield individual item Performance and Satisfaction scores. The activity item scores are then averaged to give the total Performance and Satisfaction scores that are used to make comparisons between the pre-test and post-test assessment occasions.

The scope of this study did not allow a factor analytic review nor Rasch analysis of the structural fidelity of the COPM. Analyses were limited to the study of inter-item relationships to shed some light on the dimensionality of the Importance ratings, and Performance and Satisfaction with performance subscale scores.

Importance Rating As a Weighting System

In the COPM, the activities that clients perceive as being difficult to perform are not necessarily the most important to their daily life. By the same token, clients may not find activities which are important to them difficult to perform. Therefore, the convergent inter-item relationships among the Importance ratings (weightings) of the activity items were predicted to be comparatively low.

Following the procedures of the COPM, five activities with the highest Importance ratings are selected to be the activity items for further rating.

For this part of the analysis, the activity items of each client were ranked with respect to their Importance ratings and assigned to Item 1 to 5 accordingly. Item 1 is the activity item with the highest Importance rating whereas Item 5 is the lowest. Table 5.18 summarizes the mean Importance ratings and their standard deviations. The analysis found the mean Importance ratings of the five activity items ranged from 7.6 (SD=2.3) to 9.8 (SD=0.7) (Table 5.18).

Table 5.18

Inter-item Correlation of Importance Ratings of COPM's Activity Items

Item Importance Ratings (Pearson's *r*)				
Act. Item 1	Act. Item 2	Act. Item 3	Act. Item 4	Act. Item 5
M=9.8	M=9.2	M=8.8	M=8.1	M=7.6
SD=0.7	SD=1.7	SD=1.8	SD=2.2	SD=2.3

Note. Act. Item 1 is the activity with the highest Importance rating and Act. Item 5 with the lowest for each client.

The elimination of activities deemed less important to the clients introduced biases into the scoring system in two ways. First, the weighting system in the COPM was most likely rendered meaningless since most of the activity items had a very limited range of Importance ratings. Second, the construct of "importance" in clients' daily life was skewed toward the low end of the scale. The truthfulness and meaningfulness of the Importance ratings as a weighting system in the COPM, and the subsequent computation Performance and Satisfaction scores are called into question.

The Performance Scale

Previous analyses showed a clear conceptualization of the construct of "performance" among the clients. The Performance subscale was found to be meaningful in reflecting different performance levels among the clients.

The Performance score of each activity item was computed by multiplying the Performance rating by the Importance rating. Like the Importance ratings, the structural relationships among the item Performance scores should be relatively weak due to the differences in the nature of activities selected by clients, their subsequent Importance ratings, and clients' perceived abilities to perform those activities.

In this part of analysis, activity items were arranged into Act. Item 1 to 5 according to the sequences that appeared on the COPM test protocol (Table 5.19). In other words, the sequence of Act. Item 1 to 5 is the sequence with which clients assigned their perceived Performance and Satisfaction with Performance ratings. Table 5.19 presents the inter-item correlations among the item performance scores of the initial and pre-discharge assessments. The figures along the diagonal were the correlations of the item performance scores between the initial and pre-discharge assessments.

Table 5.19

Inter-item Correlations of Performance Scores of COPM's Activity Items in Initial and Pre-discharge Assessments

	Item Performance Scores (Pearson's *r*)				
	Act. Item 1 M=29.1 SD=21.5 (M=70.2) (SD=24.3)	Act. Item 2 M=29.5 SD=25.0 (M=67.9) (SD=24.3)	Act. Item 3 M=24.8 SD=22.8 (M=60.1) (SD=31.1)	Act. Item 4 M=18.5 SD=15.2 (M=50.0) (SD=29.1)	Act. Item 5 M=20.0 SD=14.7 (M=44.5) (SD=29.1)
Act. Item 1	**.22**	.42*	.20	-.04	.01
Act. Item 2	.20	**.07**	.55**	.12	.15
Act. Item 3	.27	.47*	**.04**	.23	.26
Act. Item 4	.31	.54**	.60**	**.10**	.52*
Act. Item 5	.13	.23	.32	.70**	**.42***

Note. * $p \leq .05$ ** $p \leq .001$

Act. Item 1 is the first rated activity item and Act. Item 5 is the last. Correlation coefficients above (below) the diagonal are for the initial assessment (pre-discharge assessment). Means and standard deviations in parentheses represent the pre-discharge assessment.

Initial performance score. The mean Performance ratings of each of the five activity items across clients in the initial assessment occasion ranged from 2.3 (SD=1.7) to 3.3 (SD=2.7) (Table 5.22). The low means indicated that clients generally perceived their performance in the selected activities as "poor". This could be due to the fact that clients were generally hindered by their dysfunction in the performance components at the early stage of the rehabilitation programs which was reflected in their low ratings. The mean Performance scores of each of the five activity items, after being multiplied by the Importance ratings, varied from 18.5 (SD=15.2) to 29.5 (SD=25.0) on the initial assessment (Table 5.19).

Table 5.19 shows the correlations among the item Performance scores of the five activity items at the initial assessment (correlations above the diagonal line) and the pre-discharge assessment (correlations below the diagonal line). For the initial assessment, there was a tendency for adjacent items to have higher correlations than the non-adjacent items. For example, items 2 and 3 yielded a correlation coefficient of 0.55 but for items 2 and 5, the correlation was only 0.15.

Pre-discharge performance scores. It was speculated that the relationships among the item Performance scores were likely to remain moderate in the pre-discharge assessment. Furthermore, the mean Performance ratings were predicted to be higher and more varied than that of the pre-test because of the treatment programs that the clients received (as interventions). The pre-discharge mean Performance ratings of activity items revealed a wider range from 5.5 (SD=3.1) to

7.4 (SD=2.4) (Table 5.24). The pre-discharge Performance scores varied from 44.5 (SD=29.1) to 70.2 (SD=24.3) (Table 5.19). The patterns and sizes of the inter-item correlations were similar to that of the initial assessment with a mixture of high (*r*=.47 to .70) and low (*r*=.13 to .32) values (Table 5.19).

The patterns of the inter-item correlations of both the initial and pre-discharge Performance scores suggested a "carry-over" effect in which the performance rating assigned to one item influenced the rating assigned to the next activity item. For example, in the initial assessment occasion, items 2 and 3 yielded a correlation coefficient of .55 whereas items 2 and 5 was .15; in the pre-discharge assessment occasion, items 2 and 3 yielded a correlation coefficient of .47 whereas items 2 and 5 was .23. However, the effects of the first Performance rating assignment diminished as the clients proceeded with ratings to the next few activity items. Other than this, no systematic pattern was revealed which further substantiates the notion that individual clients perceived their levels of functioning differently for different activity items.

Initial and pre-discharge item performance scores. Low correlations of the item performance scores were revealed between the initial and pre-discharge assessments (Table 5.19). This further substantiated the notion that clients' responses on the COPM on different assessment occasions were different. It also partially explained the low stability estimate yielded for the Performance subscale.

The Satisfaction Scale

In the COPM, after the clients rate their current performance on the activity items, they are asked to assign Satisfaction ratings to their perceived performance with a 10-point rating scale. The item Satisfaction scores are computed by multiplying the Satisfaction rating by the Importance rating. Like the Performance Scale, the relationships among the Satisfaction scores should be relatively weak due to the differences in perceived performance and subsequent satisfaction with that performance among individual clients. Furthermore, results of previous analyses of substantive validity of the COPM have indicated that clients interpreted the concept of "Satisfaction" and its ratings differently. This further weakens the relationships among the Satisfaction scores of the activity items. Table 5.20 presents the inter-item correlations among the item satisfaction scores of the initial and pre-discharge assessments. The figures along the diagonal were the correlations of the item satisfaction scores between the initial and pre-discharge assessment.

Table 5.20

Inter-item Correlation of Satisfaction Scores of COPM's Activity Items in Initial and Pre-discharge Assessments

	Item Satisfaction Scores (Pearson's *r*)				
	Act. Item 1 M=31.4 SD=23.4 (M=73.7) (SD=22.6)	Act. Item 2 M=22.9 SD=22.8 (M=61.3) (SD=29.6)	Act. Item 3 M=25.4 SD=25.7 (M=59.3) (SD=32.0)	Act. Item 4 M=20.8 SD=18.5 (M=51.3) (SD=30.3)	Act. Item 5 M=23.1 SD=20.5 (M=45.2) (SD=30.5)
Act. Item 1	**.09**	.64**	.42*	.37*	.41*
Act. Item 2	.28	**.12**	.44*	.41*	.51*
Act. Item 3	.32	.49*	**.18**	.68**	.28
Act. Item 4	.44*	.61**	.71**	**.08**	.44*
Act. Item 5	.26	.41*	.39*	.72**	**.34**

Note. * $p \leq .05$ ** $p \leq .001$

Act. Item 1 is the first rated activity item and Act. Item 5 is the last. Correlation coefficients above (below) the diagonal are for the initial assessment (pre-discharge assessment). Means and standard deviations in parentheses represent the pre-discharge assessment.

Initial Satisfaction Scores. The initial mean Satisfaction ratings of the five activity items fell within a restricted range from 2.5 (SD=2.4) to 3.4 (SD=2.4). The small variations of the Satisfaction ratings indicate that clients were generally not satisfied with their performance in the activities at the early stage of the rehabilitation programs. The mean Satisfaction scores of each of the five activity items, after being multiplied by the Importance ratings, varied from 20.8 (SD=18.5) to 31.4 (SD=23.4). Table 5.20 shows the inter-item correlations of the Satisfaction scores, indicating low to moderate relationships among the scores (*r* ranged from .37 to .68).

The two plausible explanations to these unexpectedly high inter-relationships are that: 1) clients were generally dissatisfied with their own performance (which is not uncommon in the initial phase of the rehabilitation process); 2) the meaning of "Satisfaction" tended to be interpreted differently by individual clients (as mentioned in the previous sections). Moreover, "satisfaction" could not be easily delineated with respect to different activities. Rather it was an overall belief or perception which was likely generalized to all concerns of the client. The following analysis of the pre-discharge item Satisfaction scores will shed some light on the validity of these two arguments.

Pre-discharge satisfaction scores. The pattern of inter-item correlations between the pre-discharge Satisfaction scores was also similar to that of the initial assessment (Table 5.20). The mean Satisfaction ratings of individual items presented a much wider range of 5.6 (SD =3.3) to 7.8 (SD = 2.0) than that of the initial assessment (Table 5.24). The mean Satisfaction scores of each activity items varied from 45.2 (SD=30.5) to 73.7 (SD=22.6) which is consistent with the predictions made. The high inter-item correlations (*r* ranged from .41 to .72) further substantiate that "satisfaction" is likely to be an overall belief or perception which is easily generalized to all concerns of the clients. Higher item Satisfaction scores indicate that clients were generally more satisfied with their own performance in the pre-discharge phase than in the initial phase of the rehabilitation program.

Initial and pre-discharge item satisfaction scores. Low correlations of the item satisfaction scores were revealed between the initial and pre-discharge assessments (Table 5.20). This further substantiated the notion that clients' responses on the COPM on different assessment occasions were different. It also partially explained the low stability estimate yielded for the Satisfaction with performance subscale.

Constructs of Performance and Satisfaction

In Section B of this chapter, protocol analysis of the interviews with the clients in the retest occasion revealed that clients' were unclear if not confused with the concept of "satisfaction." The constructs of perceived performance (Performance rating) and satisfaction with performance (Satisfaction rating) were thought to be inter-changeable. A correlational matrix showing the relationships between the clients' pretest Performance and Satisfaction scores on the five activity items across clients selected items is shown in Table 5.21.

Table 5.21

Correlations of Initial Item Performance and Satisfaction Scores on the COPM

Item Performance Scores	Item Satisfaction Scores (Pearson's *r*)				
	SATSC1	SATSC2	SATSC3	SATSC4	SATSC5
PERSC1	**.60****	.05	.11	.11	.01
PERSC2	.30	**.37***	.47*	.48*	.16
PERSC3	.14	.13	**.66****	.47*	.19
PERSC4	-.02	.27	.04	**.50***	.27
PERSC5	.10	.25	.21	.43*	**.73****

Note. * $p \leq .05$ ** $p \leq .001$

In the initial assessment, consistent patterns of relationships were revealed between the item Performance and Satisfaction ratings. Although the correlations were not exceedingly high, significant correlations (*r* ranged from .37 to .73) were revealed between the Performance and Satisfaction scores of the activity items (Table 5.21). The results suggested that when clients rated their performance as "poor", they would likely rate their satisfaction with that performance as "low". When the performance was rated as "good", the Satisfaction ratings were likely to be "high". The mean Performance and Satisfaction ratings of the clients on the COPM's activity items revealed similarities between the two sets of ratings (Table 5.22)

Table 5.22

Initial Mean Performance and Satisfaction Ratings of Activity Items

Item Number	Mean Performance Rating		Mean Satisfaction Rating	
	Mean	SD	Mean	SD
Act. Item 1	3.1	2.2	3.4	2.4
Act. Item 2	3.3	2.7	2.5	2.4
Act. Item 3	2.8	2.4	2.8	2.6
Act. Item 4	2.3	1.7	2.7	2.3
Act. Item 5	2.5	2.0	2.9	2.6

In the pre-discharge assessment, positive relationships between performance and satisfaction on the same item were again obvious. Table 5.23 presents the correlation matrix of the item Performance and Satisfaction scores. The significant high correlations (*r* ranged from .86 to .98) between the Performance and

Satisfaction scores on the same item suggest that it was highly likely for clients to rate their satisfaction as high when their performance was good, i.e., perceived progress towards independence. Like the initial assessment, the mean Performance and Satisfaction ratings of the clients on the COPM revealed few differences between the two sets of scores (Table 5.24).

Table 5.23

Correlations of Pre-discharge Item Performance and Satisfaction Scores on the COPM

Item Performance Scores	Item Satisfaction Scores (Pearson's *r*)				
	SATSC1	SATSC2	SATSC3	SATSC4	SATSC5
PERSC1	**.86****	.11	.20	.31	.19
PERSC2	.31	**.86****	.41*	.52**	.22
PERSC3	.35*	.47*	**.93****	.66**	.33
PERSC4	.47*	.64**	.64**	**.93****	.73**
PERSC5	.19	.44*	.38*	.69**	**.98****

Note. * $p \leq .05$ ** $p \leq .001$

Table 5.24

Pre-discharge Mean Performance and Satisfaction Ratings of Each Activity Item

Item Number	Mean Performance Rating		Mean Satisfaction Rating	
	Mean	SD	Mean	SD
Act. Item 1	7.4	2.4	7.8	2.0
Act. Item 2	7.4	2.2	6.6	2.9
Act. Item 3	6.6	2.9	6.4	3.0
Act. Item 4	6.0	2.9	6.1	2.9
Act. Item 5	5.5	3.1	5.6	3.3

Another method used to substantiate the relationships between the Performance and Satisfaction ratings was to correlate the change in Performance ratings and the Satisfaction ratings by subtracting the ratings of the pre-discharge assessment from their initial assessment counterparts for each activity item. The strong correlations between the two sets of change in ratings provide further insights into the relationships of the constructs of performance and satisfaction with the performance as shown below.

Positive and significant correlations were revealed between the change in Performance and Satisfaction ratings for all the activity items (Table 5.25). Correlation coefficients *r* along the diagonal of Table 5.25 ranged from .64 to .89. This further suggests that the constructs of performance and satisfaction with the performance are closely related to one another. Their relationships are consistent with those speculated from the model of occupational performance and the findings in the protocol analysis of this study.

Table 5.25

Correlation of Change in Performance and Satisfaction Ratings on the COPM

Change in Performance Ratings	Change in Satisfaction Ratings (Pearson's r)				
	CHSAT1	CHSAT2	CHSAT3	CHSAT4	CHSAT5
CHPER1	.64**	-.06	.15	.20	.12
CHPER2	.19	.64**	.50*	.38*	.07
CHPER3	.15	.50*	.82**	.58**	.33*
CHPER4	.23	.58*	.64**	.89**	.57**
CHPER5	.14	.31	.60**	.69**	.89**

Note. * $p \leq .05$ ** $p \leq .001$

Strong and positive relationships were also revealed between the Performance and Satisfaction subscales in both initial and pre-discharge assessment occasions (Table 5.26). Pearson's *r* between the two subscales was .77 ($p < .001$) for the initial assessment and .94 ($p < .001$) for the pre-discharge assessment. The high correlations between the two COPM subscale scores yielded in the same assessment occasion further substantiate the convergence of the constructs of performance and satisfaction with performance as revealed in the previous analyses. Low correlations were found between the subscale scores yielded in different assessment occasions, with *r* ranging from .17 to .28.

Table 5.26

Correlation of COPM Performance and Satisfaction Subscale Scores in Initial and Pre-discharge Assessments

	COPM Subscale Scores (Pearson's *r*)			
	Initial Performance Scores	Initial Satisfaction Scores	Pre-discharge Performance Scores	Pre-discharge Satisfaction Scores
Initial Performance Scores	1.00	.77**	.17	.15
Initial Satisfaction Scores		1.00	.25	.28
Pre-discharge Performance Scores			1.00	.94**
Pre-discharge Satisfaction Scores				1.00

Note. * $p \leq .05$ ** $p \leq .001$

Conclusion

Results of the analyses suggest that clients' ratings of their Performance and Satisfaction subscales are closely related to one another. The literature review and the protocol analyses conducted in this study have demonstrated the commonalities and interrelationships between these two constructs. Statistical findings also support the notion of the common variance and high correlations between the two subscales in the COPM. This leads to a question about the need for incorporating two highly related and dependent measurement scales in one instrument. From a psychometric point of view, the two subscale scores do not provide independent contributions towards better discrimination and prediction of clients' occupational performance. From a clinical point of view, the use of the Performance as well as Satisfaction subscales provides little additional essential information than the use of either alone.

The analyses revealed the problems with the dimensionality of the Performance and Satisfaction with performance subscales of the COPM. Consequently, the clinical interpretations of the subscale scores in the COPM are called into question. The usefulness of the comparisons between the initial and pre-discharge test scores of individual clients are also challenged. Hence, the psychometric properties of the COPM as a clinical outcome measure are at best weak.

Expert Panel Review of Structural-Related Evidence : A Qualitative Analysis

A different perspective for evaluating the structural validity of the COPM scoring system used data extracted from the Structural Review portion of the expert panel review questionnaire. The expert panel structural review was made up of eight items that were designed to gather opinions from the panel members on the meaningfulness of the scoring model as a means of reflecting clients' occupational performance. Panel members used a 5-point Likert scale to evaluate the scoring model. The model includes the use of the 10-point rating scales, and the computation methods of the item and total Performance and Satisfaction scores in the COPM. A sample item from the questionnaire is as follows:

> To what degree does the computation of the Performance scores (multiply Importance rating by Performance rating) for each identified activity (item) reflect the structural relations between a person's occupational performance and performance components? Rating:
>
> | 1 | 2 | 3 | 4 | 5 |
> |---|---|---|---|---|
> | Poor | Fair | Good | Very Good | Excellent |
>
> Suggestions for changes, if any:
>
> ____
>
> ____

The above item was constructed to solicit the panel members' evaluation of the extent to which the scoring model meaningfully reflected the occupational performance of the clients on the COPM. In responding to this item, the panel members were expected to interpret the COPM Performance scores (Importance rating x Performance rating) in terms of clients' clinical conditions and their performance of activities (such as "dressing upper limb garment"). The open-ended portion of the item was used to seek panel members' suggestions for future changes.

Use of 10-point Rating Scale in COPM

The first three items (2.1 - 2.3) in the structural review reflected the relevance of using the 10-point rating scale to quantify the constructs of Importance, Self-Perceived Performance, and Satisfaction with Performance of clients in the COPM. The mean rating in this area was 3.2 (3.5 when the 9th rater's score was removed). This suggests that on average raters regarded the use of the 10-point rating scale as "good" (Table 5.27) to quantify the importance, performance, and satisfaction of clients. All three scales were found to have similar ratings (3.6, 3.4, and 3.6 for item 2.1, 2.2, and 2.3 respectively when the 9th rater's scores were removed).

Table 5.27

Structural Fidelity of Using the 10-point Rating Scale in the COPM

Item No. Rating Scale	Raters									Item Mean Scores
	1	2	3	4	5	6	7	8	9	
2.1 Importance Rating	4	4	4	3	3	3	3	5	1	3.3
2.2 Performance Rating	4	4	2	3	3	3	3	5	1	3.1
2.3 Satisfaction Rating	4	4	3	3	3	3	4	5	1	3.3
Rater Mean Scores	4.0	4.0	3.0	3.0	3.0	3.0	3.3	5.0	1.0	3.2

In reviewing the open-ended comments, the majority of the raters were satisfied with the use of the 10-point rating scale. Criticisms of the 10-point scale were made against the arbitrary anchor points. One rater offered this criticism:

> "I think this type of scale is as good as any for this purpose except I think there should be a zero value, e.g. 0 - 10 for all scales. Then everything will be okay."

Other criticisms included the lack of a neutral point and the use of only integer ratings.

Computation of Scores on COPM

The remainder of the items in the structural review (2.4 - 2.8) were intended to evaluate the extent to which the computation of the item and total Performance and Satisfaction scores on the COPM reflect the structural relations between occupational performance and performance components. The mean rating in this area was 2.4 (2.7 when the 9th rater was removed) suggesting that the raters regarded the computational methods used in the COPM as "fair" to "good" in quantifying clients' occupational performance (Table 5.28).

A more in-depth analysis of the raters' ratings reveals a clear disparity within the group of raters. Table 5.28 shows the ratings for the five activity items. As a group, the ratings can be generally divided into high (3.0 or above) (*n*=4) and low groups (2.2 or below) (*n*=5).

Table 5.28

Structural Fidelity of Computation of Item and Total Scores on the COPM

Item No. Computation of Scores	Raters									Item Mean Scores
	1	2	3	4	5	6	7	8	9	
2.4 Item Performance Score	1	4	3	3	3	1	1	4	1	2.3
2.5 Item Satisfaction Score	1	4	3	3	3	1	1	4	1	2.3
2.6 Total Performance Score	2	4	2	3	3	2	1	4	1	2.4
2.7 Total Satisfaction Score	2	4	2	3	3	2	1	4	1	2.4
2.8 Change Scores	2	3	1	3	5	3	3	2	1	2.6
Rater Mean Scores	1.6	3.8	2.2	3.0	3.4	1.8	1.4	3.6	1.0	2.4

The raters in the high rating group did not provide comments. However, comments provided by three raters in the low rating group shed light on their dissatisfaction with the quantification methods used in the COPM. They all commented that the computation of the Performance and Satisfaction scores was not meaningful. They did not know how to interpret the numbers and relate the numbers to the occupational performance of the clients. Instead, they reflected that the scores may have provided some indications of clients' orientation to the environment and perception of how well they performed. Some examples of their comments are,

> "I don't know what multiplying the scores (Importance x Performance) accomplishes...Averaging seems to cloud the issue."

Another rater wrote,

> "The scores seem meaningless to me because there is no scale to compare it to. That is, the 10-point scale can only give a certain degree of qualitative data."

A third rater added,

> "Do we need to quantify them? What sense can we make out of it because clients' performance can be out of very different constructs."

Four raters wrote comments on the item asking how well the change scores (both Performance and Satisfaction) on the COPM quantified the changes in clients' occupational performance. They queried the meaningfulness of subtracting the pre-discharge COPM Performance and Satisfaction scores from those obtained in the initial assessment. Furthermore, two raters criticized the lack of a norm for facilitating clinical comparisons and decision making.

Two raters commented on the psychometric aspects of deriving the Performance and Satisfaction change scores. For example, one rater identified a major flaw in the COPM in the assumption that the importance ratings of activities remain unchanged from initial to pre-discharge occasions. In fact, it is not uncommon for clients to change their life roles throughout the rehabilitation processes based on their residual capabilities. For instance, the role of a worker could be changed to that of a home-maker due to one's inability to perform work tasks which are too demanding for a post-stroke client. Another rater criticized the inappropriateness of limiting the number of activities selected to five. In fact, a lot of clients that this particular rater encountered had ten or more problems with their occupational performance which were of equal importance. However, based on the testing rules, other comparatively less important activities have to be eliminated from the ratings on the COPM. Thus, the meaningfulness of how the scores in the COPM reflect the changes within clients' occupational performance was challenged.

Conclusion

The use of the 10-point rating scale has been recognized as appropriate in its quantification of importance, perceived performance, and satisfaction of clients. However, the validity of using a 1 to 10 point scale was challenged. Instead, as recommended by one rater, a 0 to 10 point scale is more meaningful as the rating scale of the COPM.

In general, the computation of item and total Performance and Satisfaction scores on the COPM was criticized as "meaningless" in terms of its clinical interpretations and decisions. Clinical norms may be difficult to develop because of the differences in the life roles and importance of individual activities among different clients. Other psychometric problems such as the changes in clients' concerns with their problems between the initial and pre-discharge assessment occasions were also regarded as critical to the meaningfulness of quantifying the Performance and Satisfaction scores in the COPM to represent clients' occupational performance and that changes throughout the rehabilitation process.

Section D - Criterion-Related Evidence

In the last Section of this chapter, the criterion-related evidence of construct validity of the COPM is presented. As mentioned in Chapter IV, the small sample size of this study limited the types of analysis that could be conducted in establishing convergent and discriminant evidence between the COPM and other "criterion" measures. Hence, inferences made in this part of the analysis are based on simple correlational statistical methods. The three "criterion" measures are the Klein-Bell ADL Scale, Satisfaction with Performance Scaled Questionnaire, and the Functional Independence Measure. Each of the instruments is similar to one subscale of the COPM either in the test trait (or construct) or testing method. The ways in which each of the instruments establishes convergent or discriminant evidence with the COPM are presented in Tables 5.29, 5.31, and 5.33. The framework of analysis in this part is based on the multi-trait multi-method model as suggested by Campbell and Fiske (1959).

COPM and Klein-Bell ADL Scale

The COPM is a self-report clinical outcome measure that assesses a client's performance in activities selected by the client at different stages of occupational therapy intervention. A client's performances in the selected activities is termed "occupational performance". The Klein-Bell ADL Scale is a therapist's assessment of a client's performance on a set of well-defined self-care activities. Although the test does not specify its use as a repeated measure, it has been commonly used by clinicians to measure a client's improvement in receiving occupational therapy.

Table 5.29 summarizes the expected relationships between the COPM and the Klein-Bell ADL Scale in terms of traits and methods. Basically, the Klein-Bell ADL and the COPM Performance subscale assess similar traits but with different methods. The correlations between the COPM Performance subscale score and the five Klein-Bell ADL subscale scores (the communication subscale of the Klein-Bell ADL scale was not included due to its zero variance) were predicted to be moderate. In contrast, the Klein-Bell subscales are different from the COPM Satisfaction subscale in both the trait and method. Consequently, the correlations between these subscale scores were expected to be low.

Table 5.29

Expected Convergent and Discriminant Validity between COPM and KB-ADL Scale

COPM Scores	Klein-Bell ADL Scale Scores
Total Performance	Similar Trait, Different Method
Total Satisfaction	Different Trait, Different Method
Change in Performance	Similar Trait, Different Method
Change in Satisfaction	Different Trait, Different Method

Initial Assessment Occasion

The correlation coefficients between the COPM and Klein-Bell ADL Scale scores for the initial assessment are presented in Table 5.30. As shown, the patterns of relationships that were predicted were not found. Convergent validity was not observed between the COPM Performance subscale and the Klein-Bell activity subscales. In fact, some of the correlation coefficients were virtually zero or even negative in value. The same pattern of relationships was observed in the COPM Change in Performance score with the Klein-Bell subscales.

Since there was no support for convergent validity, discriminant validity between the COPM Satisfaction subscale and Klein-Bell subscales became difficult to interpret. The correlations between them were weak and mostly negative. The relationships between the Change in Satisfaction and Klein-Bell subscale were also weak but positive. Only one of the correlation coefficients was significant which was likely attributable to the instability of the COPM in assessing clients' occupational performance.

Table 5.30

Convergent and Discriminant Validity between COPM and KB-ADL Scale for Initial and Pre-discharge Assessment

COPM Total and Change Scores	Klein-Bell ADL Scale Scores (Pearson's *r*)				
	Dressing	Mobility	Bathing	Eating	Elimination
Total Performance	.08 (.21)	-.10 (.16)	-.18 (-.05)	-.20 (.03)	.04 (.33*)
Change in Performance	.11 (.15)	.03 (.14)	.00 (-.05)	.01 (.10)	.17 (.25)
Total Satisfaction	-.15 (.13)	-.18 (.14)	-.32* (-.05)	-.12 (-.07)	.03 (.26)
Change in Satisfaction	.10 (.09)	.12 (.27)	.23 (.02)	-.23 (.02)	.19 (.21)

Note. * indicates $p \leq .05$.

Correlation coefficients in parentheses represent relationships between the two sets of scores obtained in the pre-discharge assessment.

In general, the patterns of convergent and discriminant validity were not observed. Besides the small sample size, this "misbehavior" of the results is explained by the flaws in the design of the COPM and various threats to its construct validity as presented in previous sections.

Pre-discharge Assessment Occasion

Similar to the results revealed in the initial assessment occasion, the convergent validity of the COPM Performance subscale and Klein-Bell subscales was not observed in the pre-discharge assessment scores (Table 5.30). In fact, the relationships between both the COPM Performance and Satisfaction subscales and the Klein-Bell subscales were similar. Stronger relationships were observed between the Klein-Bell Elimination subscale and the COPM subscales, particularly with the total Performance score ($r=.33$, $p<.05$). However, one piece of positive evidence was not strong enough to substantiate the convergent validity between the two instruments.

COPM and SPSQ

The SPSQ is a clinical self-report assessment of clients' satisfaction with their own performance in two areas of activities, namely, home management and social/community problem solving. Thus, the SPSQ assesses a trait similar to the Satisfaction subscale of COPM with a similar method (Table 5.31). However, the diversity of the activities selected by clients in the COPM does not allow the prediction of which subscales of the SPSQ will have a stronger relationship with the subscales of the COPM. As a consequence, moderate correlations were predicted between the two sets of scores. In contrast, while the method is similar, discriminant validity was predicted to be established between the COPM Performance subscale and the SPSQ subscales because of the different traits assessed by each. The correlations hence between them should be weak.

Table 5.31

Expected Convergent and Discriminant Validity between COPM and SPSQ

COPM Total and Change Scores	Satisfaction with Performance Scaled Questionnaire Scores: Home Management Subscale	Satisfaction with Performance Scaled Questionnaire Scores: Social/Community Problem Solving Subscale
Total Performance	Different Trait, Similar Method	
Change in Performance	Different Trait, Similar Method	
Total Satisfaction	Similar Trait, Similar Method	
Change in Satisfaction	Similar Trait, Similar Method	

Initial Assessment Occasion

The correlation coefficients of the initial subscale scores between the COPM and SPSQ are presented in Table 5.32. The findings do not reveal the predicted pattern of relationships. Convergent validity was not observed between the COPM Satisfaction subscale and SPSQ subscales. The correlation coefficients were virtually zero. The same lack of support for convergent validity was observed in the COPM Change in Satisfaction score and the SPSQ subscales. The relationships between the COPM Performance subscale and the Change in Performance scores, and the SPSQ subscale scores were weak. The low short term stability of the instrument in measuring clients' occupational performance was speculated to be a major contributory factor to the insignificance of relationships.

In general, the patterns of convergent and discriminant validity were not observed between the COPM and SPSQ. As with the Klein-Bell ADL Scale, the mis-behavior of the results was explained by the flaws in the design of the COPM and various threats to its construct validity.

Table 5.32

Convergent and Discriminant Validity between the COPM and SPSQ for Initial and Pre-discharge Assessments

COPM Total and Change Scores	Satisfaction with Performance Scaled Questionnaire Scores (Pearson's *r*)	
	Home Management Subscale	Social/Community Problem Solving Subscale
Total Performance	.18 (.31*)	.02 (.39*)
Change in Performance	.02 (.25)	-.02 (.39*)
Total Satisfaction	-.02 (.22)	-.13 (.36*)
Change in Satisfaction	.03 (.27)	-.01 (.40*)

Note. * indicates $p \leq .05$.

Correlation coefficients in parentheses represent relationships between the two sets of scores obtained in pre-discharge assessment.

Pre-discharge Assessment Occasion

Stronger relationships between the pre-discharge scores on the COPM and the SPSQ were revealed based on their significant and higher correlation coefficients than those of the initial assessment (Table 5.32). The SPSQ social/community subscale had significant and positive relationships (*r* ranged between .36 and .40) with the two subscales scores and the Change scores of the COPM. However, no differentiated convergent and discriminative validity was shown among them. The SPSQ home management subscale had weaker and insignificant relationships with the COPM scores, and so convergent and discriminant validity were not evident.

The patterns of relationships among the pre-discharge COPM and SPSQ scores show that measurement of the Performance and Satisfaction constructs in the COPM are not independent. Some commonalities with the social/community and, to a lesser extent, home management subscales were shown. The findings in this part are consistent with those in the earlier substantive validity section of this chapter.

COPM and Functional Independence Measure

The Functional Independence Measure incorporates two subscales, the motor and cognitive. The instrument claims to measure the functional status, burden of care, and level of disability of clients. The construction of the instrument uses a minimum criterion concept which means the number of items in both subscales are those most critical to the discrimination of clients with independent versus dependent functioning. The testing method used in the FIM is a performance rating administered by clinicians. Convergent validity was predicted to be established between the COPM Performance subscale and FIM motor subscale as they had similar traits but different methods of measuring clients' functioning (Table 5.33). Discriminant validity was predicted between the COPM Satisfaction subscale with the FIM motor and cognitive subscales because of their differences in both trait and method. Similar patterns were expected in the COPM Change in Performance and Change in Satisfaction scores with the FIM subscales.

Table 5.33

Expected Convergent and Discriminant Validity between COPM and FIM

	Functional Independence Measure Subscale Scores	
COPM Total and Change Scores	Motor Subscale	Cognitive Subscale
Total Performance	Similar Trait Different Method	Different Trait Different Method
Change in Performance	Similar Trait Different Method	Different Trait Different Method
Total Satisfaction	Different Trait Different Method	Different Trait Different Method
Change in Satisfaction	Different Trait Different Method	Different Trait Different Method

Initial Assessment Occasion

The relationships between the COPM Performance subscale and FIM motor subscale were not marked and significant (Table 5.34). Similarly, the relationships between the FIM motor and cognitive subscales with the rest of the COPM subscale and Change scores were diverse and insignificant. Predicting of convergent and discriminant validity between the COPM and FIM failed to occur.

Table 5.34

Convergent and Discriminant Validity between COPM and FIM for Initial and Pre-discharge Assessment

COPM Total and Change Scores	Functional Independence Measure Subscale Scores (Pearson's Correlation Coefficient) Motor Subscale	Cognitive Subscale
Total Performance	-.03 (.32*)	-.17 (.20)
Change in Performance	.15 (.38*)	.22 (.19)
Total Satisfaction	-.14 (.26)	-.16 (.14)
Change in Satisfaction	.16 (.36*)	.14 (.16)

Note. * indicates $p \leq .05$.

Correlation coefficients in parentheses represent relationships between the two sets of scores obtained in the pre-discharge assessment.

Like the Klein-Bell ADL and SPSQ, the negative and insignificant results were likely attributable to the low short term stability of the COPM and the threats of its construct validity.

Pre-discharge Assessment Occasion

Results obtained in the pre-discharge assessment revealed clearer patterns of relationships between the COPM and FIM subscales (Table 5.34). Convergent validity was found between the COPM Performance subscale scores and FIM motor subscale scores as predicted. Correlation coefficients suggested significant but weak positive relationships between the FIM motor subscale score, and the COPM Performance subscale and the Change scores ($r=.32$ and .38 respectively, $p<.05$). Discriminant validity was demonstrated between the FIM cognitive subscale and the COPM Performance subscale and Change in Performance. Correlation coefficients reflected weaker positive relationships among them ($r=.20$ and .19 respectively).

Discriminative validity was not well established between the COPM Satisfaction subscale and the FIM motor subscale. The Pearson's r between the COPM Satisfaction and the FIM motor subscale scores was .26 ($p>.05$). However, a higher and significant correlation ($r=.36$, $p<.05$) was found between the COPM Change in Satisfaction scores and the FIM motor subscale scores.

Conclusion

The multi-trait and multi-method strategy has been regarded as a useful method in establishing criterion-related evidence of clinical assessment tools (Campbell & Fiske, 1959; Rothman, Hedrick, & Inui, 1989). The experience gained in this section suggests that the usefulness of this method is compromised when the construct specificity and, to a certain extent, the reliability of the instrument being validated have not yet been ascertained. In the COPM validation process, evidence gathered in the content, substantive and structural reviews have all revealed threats to construct validity of the instrument. As a consequence, the failure to establish significant convergent and discriminant validity between the COPM and its criterion was not surprising.

The multi-trait multi-method analyses in this section prompt the following observations:

As predicted:

1) Significant and moderate relationships were found between the FIM Motor subscale, and the Performance subscale and Change in Performance of the COPM at the pre-discharge occasion.

2) Significant and moderate relationships were found between the SPSQ Social/Community Problem Solving subscale, and the Satisfaction subscale and Change in Satisfaction of the COPM at pre-discharge occasion.

Not as predicted:

1) The COPM Performance and Satisfaction subscales of the COPM were found to be closely related to one another which made it difficult to discern convergent and discriminant validity between these two subscales and the "criterion" instruments.

2) Relationships established between the COPM subscale scores and their "criterion" subscales scores were weaker in the initial assessment results than in the pre-discharge results which revealed significant relationships. This wide disparity of results between the initial and pre-discharge assessments supports the argument made earlier that measurement errors are associated with different assessment occasions. Findings also support the argument that the testing effects influenced clients' responses on the COPM in the pre-discharge assessment occasion especially for those who participated in the retest of the study. The serious measurement errors which occured in the initial assessment were predicted to yield exceedingly low short term stability estimates.

3) Correlations between the COPM and the three "criterion" instruments have shown that the effects of "method" were not as significant as the effects of "trait". This was illustrated in the comparisons of the strengths of correlations between the SPSQ and FIM with the COPM. In principle, the SPSQ shared similar traits and methods with the COPM Satisfaction

subscale, and therefore, stronger relationships were expected there than with the FIM which shares traits but not method with the COPM Performance subscale. However, results did not reveal these predicted patterns. Instead, findings of this study were found to converge with those of other studies on the effects of different methods of clinical assessment. Elam, Graney, Beaver, Derwi, Applegate and Miller (1991) concluded that methods of clients' self-report and clinicians' judgment were valid in reflecting the actual performance of the clients. Results in this study tend to support Myers, Holliday, Harvey, and Hutchison's (1993) notion that clinical findings gathered using various methods do not differ in their qualities, instead, they serve different purposes and assist clinicians to reach different clinical decisions.

CHAPTER VI

OTHER PSYCHOMETRIC PROPERTIES

Introduction

This chapter reports the results of the panel review's evaluation of the testing procedures and utility of the COPM. Findings are presented in five areas - the testing processes, standardization, scientific rigor, clinical utility, and control of testers' competence.

The Utility and Procedural Review

Data analysed in this chapter were taken from the Utility and Procedural review portion of the expert panel review questionnaire. Details about the procedures and questionnaire of the panel review are provided in Chapters III and V.

The Utility and Procedural review consisted of 19 items designed to gather opinions on the quality and application of the COPM as an instrument to measure outcomes of clients receiving occupational therapy in hospital setting. The 19 items were constructed based on the criteria suggested by Baum (1991), Law (1987), and Loevinger (1967). All items were close-ended with the same 5-point Likert scale used in the first two sections of the questionnaire. No open-ended questions were asked in these items of the questionnaire. However, three panel members did add written comments.

Panel members were asked to evaluate the COPM in the following five major areas: testing procedures, standardization (related to reliability issues), scientific rigor (related to validity issues), clinical utility, and the control of testers' competence. The results of the evaluations are presented and analyzed with respect to these five different areas.

Evaluation of Testing Procedures of COPM

Panel members were asked to evaluate the testing procedures of the COPM in terms of the time required to administer the COPM, the use of equipment in the testing package, organization and content of the test manual, and the organization of the test format. In general, the panel members rated the COPM's overall testing procedures as "good" to "very good" (mean rating 3.7) (Table 6.1).

Panel members' ratings on the length of time required to complete the COPM varied from "fair" to "excellent". Five members rated the time required as "very good" or better. The wide variation among panel members probably reflected their personal preference on "how much" time should be spent on conducting a clinical assessment.

Eight members rated the use of testing equipment in the test package as either "very good" or "excellent". The COPM does not require a lot of testing equipment. All the tester needs to have is the test form, three rating scales, and the test manual, all of which are included in the test package.

Table 6.1

Expert Panel Review on Testing Procedures of COPM

Item No. Evaluation of Testing Procedures	Raters									Item Mean Scores
	1	2	3	4	5	6	7	8	9	
3.5 Time Required for Administration	4	2	5	2	3	4	4	5	3	3.6
3.6 Testing Equipment Required	5	4	5	5	5	4	4	5	3	4.4
3.14 Organization / Content of Test Manual	4	2	5	3	3	4	3	3	3	3.3
3.15 Organization of Test Format	4	2	4	4	3	4	3	4	3	3.4
Rater Mean Scores	4.3	2.5	4.8	3.5	3.5	4.0	3.5	4.3	3.0	3.7

On average, the panel members rated the organization and content of the Test Manual as "good" (mean rating 3.3). Members identified the main drawback as the lack of technical information about the test construction process and validation. In fact, no evidence on the psychometric properties is available in the Test Manual. Moreover, there are no guidelines on the use and interpretation of the test scores. Such information is essential in allowing clinicians to use the results of the COPM for making clinical decisions.

Eight of the panel members rated the organization of the COPM's test format as either "good" or "very good". Comments from two clinicians who had experience in administering the instrument suggested that the semi-interview format is not a good way to obtain objective results of occupational performance. The main drawback is the known "tester effect" - the subjectivity of the clinicians and familiarity of their own clients' conditions which impede the quality of the assessment results. In addition, the use of a self-reporting method to measure clients' performance of activities was challenged by some clinicians who preferred to directly observe and then rate clients' actual performance.

Evaluation of Standardization of COPM

The evaluation of test standardization of the COPM included rating the degree of clarity of the training process, clarity and preciseness of the assessment procedures, standardized client group and population, and evidence of test reliability of the COPM. In general, panel members rated the overall standardization of COPM as "fair". The item with the lowest rating - "poor" (mean rating 1.0) - was on the evidence of test reliability. The item with the highest rating of "good" was on the clarity and preciseness of the testing process (mean rating 3.2) (Table 6.2).

Table 6.2

Expert Panel Review Evaluation of Standardization of COPM

Item No. Evaluation of COPM's Assessment Standardization	Raters									Item Mean Scores
	1	2	3	4	5	6	7	8	9	
3.2 Clarity of Training Process	3	1	4	1	2	2	1	2	1	1.9
3.4 Clarity and Preciseness of Procedures	4	2	5	3	3	3	2	4	3	3.2
3.7 Standardized Population	2	1	2	1	3	1	1	4	2	1.9
3.10 Evidence of Test Reliability	1	1	1	1	1	1	1	1	1	1.0
Rater Mean Scores	2.5	1.3	3.0	1.5	2.3	1.8	1.3	2.8	1.8	2.0

Seven of the nine panel members commented that the training procedures were not very clearly presented in the Test Manual. The comments made indicated that this is particularly true for the semi-structured interview in which the clients identified activities for further evaluation of performance and satisfaction. There are no guidelines as to the amount of prompting clinicians should give to their clients.

Seven panel members rated the clarity and preciseness of the assessment procedures as at least "good". The comments made in connection with this aspect were again centered on the semi-structured interview and the lack of appropriate guidelines.

When the panel members were asked to rate the extent to which the COPM was standardized for various client populations, three indicated "fair", and four "poor". In the Test Manual, information is given only on how to use a proxy to assess clients who can not be directly assessed with the COPM. However, the procedures provided do not include procedures for clients who have language or cognitive problems.

All nine panel members evaluated the evidence of reliability of COPM as "poor". In fact, there was no information on reliability in the Test Manual. According to the Test Manual, the process of collecting evidence on reliability is "under way". However, the most recent publications of the COPM do not reveal evidence on the test reliability (Law et al., 1994; Pollock, 1993).

Evaluation of Scientific Rigor of COPM

Items in the category of scientific rigor included evaluating the method(s) used for norming and standardization, information on use and interpretation of test results (meaningfulness), interpretation of subscale and total scores, the number of references and corollary studies cited in the Test Manual, and the quality of using the COPM to measure occupational performance. In general, panel members were quite consistent in rating the scientific rigor of the COPM as "poor" or "fair" (Table 6.3).

Table 6.3

Expert Panel Review on Scientific Rigor of COPM

Item No. Evaluation of COPM's Scientific Rigor	Raters									Item Mean Scores
	1	2	3	4	5	6	7	8	9	
3.8 Norming Method	1	1	1	2	1	1	1	4	1	1.4
3.9 Meaningfulness of Test	1	1	1	1	1	1	1	4	1	1.3
3.11 Interpretation of Test Scores	1	2	3	3	3	1	2	2	1	2.0
3.18 Literature on Related Studies	1	1	4	2	2	1	1	1	1	1.6
3.19 Measure of Occupational Performance	1	2	2	2	2	2	2	2	2	1.9
Rater Mean Scores	1.0	1.4	2.2	2.0	1.8	1.2	1.4	2.6	1.2	1.6

Seven of the panel members rated the norming procedures as "poor". However, this is likely attributable to the lack of information about norming in the Test Manual. Eight panel members rated "meaningfulness of test scores" as "poor" and six rated "interpretation of test scores" as either "poor" or "fair". Similarly, these low ratings are likely due to the lack of evidence on the use and interpretation of test results either in the Test Manual or in the literature.

Eight panel members rated "literature on related studies" as either "poor" or "fair"; one member rated it "very good". Two of the panel members noted that references and corollary studies were neither cited nor reported in the Test Manual. Lastly, eight panel members rated the COPM as "fair" as a measure of occupational performance.

Evaluation of Clinical Utility of COPM

The evaluation of the clinical utility of the COPM included rating the degree to which the COPM could be used with clients across all developmental levels and disability groups, its cost effectiveness, and the degree to which the instrument was respected by other health professionals. In general, panel members rated the clinical utility of the COPM as "fair" to "good" (Table 6.4).

Table 6.4

Expert Panel Review Evaluation of Clinical Utility of COPM

Item No. Evaluation of COPM's Clinical Utility	Raters									Item Mean Scores
	1	2	3	4	5	6	7	8	9	
3.12 All Developmental Levels	2	2	2	2	3	1	4	2	4	2.4
3.13 All Disability Groups	1	3	2	1	2	1	3	2	4	2.1
3.16 Cost Effectiveness	4	2	4	3	3	2	3	4	4	3.2
3.17 Respected by Other Prof.	2	2	3	3	3	2	3	3	2	2.6
Rater Mean Scores	2.3	2.3	2.8	2.3	2.8	1.5	3.3	2.8	3.5	2.6

Most panel members gave the COPM a "fair" rating in its application to clients of different development levels; one rated this aspect as "poor". Six members rated its use with all disability group as either "poor" or "fair". Clinicians doubted whether the COPM could be applied equally to clients in different age ranges (reflecting different developmental levels), different diagnostic groups, and the use of a proxy. In the present study, the adult clients in orthopedic or stroke rehabilitation groups were screened for cognitive deficits. Data collected in the study revealed differences among individual clients such as their activity patterns, life roles, and interpretations of performance and satisfaction. If clients with more diverse developmental levels and disability groups had been included, interpretation of the COPM scores would have been even more diverse. The usefulness of the COPM for different client groups requires further research.

Three panel members thought that the cost effectiveness of using the COPM as a clinical assessment was "good"; three others indicated it was "very good". As commented by Law et al. (1994) and Pollock (1993), this is likely attributable to the relatively short time required to complete it when compared with other clinical assessments. Five panel members rated "good" and four rated "fair" the extent to which the COPM is accepted by other health professionals. However, this aspect needs more research since eight out of the nine members in the panel were occupational therapists.

Evaluation of Control of Testers' Competence

The last two items of the Utility and Procedural review concerned the evaluation of the control of testers' competence: the criteria set forth to qualify to become an examiner of the COPM and the provision of training for those examiners.

In general, the panel members commented that the guidelines set to ensure examiners' competence in administering the test as stated in the Test Manual were "good" (mean rating 2.9) (Table 6.5). However, ratings by members varied from "poor" to "very good". Similarly, wide variations were found among the ratings of panel members on the process of training competent examiners as described in the Test Manual. The mean rating was "good" with a range from "poor" to "excellent". The wide variations among panel member's ratings on these two aspects reflected individual opinions about what specific qualifications and training processes should be required in administering the COPM as a clinical assessment tool. The variability is likely due to the different clinical experiences and training among the panel members. More studies are needed to establish guidelines and standards of clinical assessments both at an instrument level and at a profession level in the future.

Table 6.5

Expert Panel Review Evaluation of Testers' Competence using COPM.

Item No. Evaluation of COPM's Testers' Competence		Raters									Item Mean Scores
		1	2	3	4	5	6	7	8	9	
3.1	Guidelines for Examiners' Training	4	2	4	3	3	3	2	4	1	2.9
3.3	Training Processes	4	1	5	3	4	1	3	4	3	3.1
Rater Mean Scores		4.0	1.5	4.5	3.0	3.5	2.0	2.5	4.0	2.0	3.0

Conclusion on Utility and Procedural Evaluation of COPM

In general, the panel members' review of the standardization, assessment procedures, and clinical utility of the COPM was "fair" or "good". The three major areas identified by the panel members as problematic were the lack of standardization of COPM's testing procedures, limitations in its clinical utility, and weak degree of scientific rigor in its construction and validation. These findings may, in part, be attributable to the fact that not much information is available on the psychometric properties of the instrument. Evidence on the usefulness and meaningfulness of using the COPM as a clinical assessment tool to measure clients' occupational performance is also limited. Despite its drawbacks at the present stage, one should regard those criticisms as an impetus for further improvement of the instrument.

CHAPTER VII

CONCLUSION AND SUGGESTIONS FOR CHANGES

Introduction

In the previous chapters, evidence concerning the construct validity of the COPM was presented. Findings revealed both strengths and weaknesses in the construct validity of the instrument. In this chapter, recommendations for possible changes of the COPM and the need for further studies will be discussed. The content of this chapter is grouped with respect to the special features, testing, and scoring procedures of the COPM. The final section of this chapter contains the conclusion of the study.

The COPM was built on the concepts of the client-centred approach and model of occupational performance (Law et al., 1994; McColl & Pranger, 1994; Pollock, 1993). In the Test Manual, the authors of the COPM described several important features of the instrument. The following characteristics were relevant to this study (Law et al., 1991, p.9-10):

The Canadian Occupational Performance Measure:

(1) is based on an explicit model of occupational therapy;

(2) encompasses the occupational performance areas of self-care, productivity, and leisure as primary outcomes;

(3) recognizes the performance components as essential to the process of occupational performance;

(4) incorporates the roles and role expectations of the clients;

(5) considers the importance of performance areas to the client;

(6) considers the client's satisfaction with present performance;

(7) can be used across all developmental levels; and

(8) can be used with all disability groups.

Features (2) to (6) are directly related to the test content of the COPM, characteristics (7) and (8) are concerned with the clinical utility of the instrument, and feature (1) is a statement about the overall test construct. These features form the basis for the following recommendations for change to improve the use and interpretation of the COPM in the measurement of occupational performance.

COPM Encompasses Three Areas of Occupational Performance

Evidence of content-related validity of the COPM was gathered from panel members' comments and from the quantitative analyses of clients' choices of activities in the COPM "Problem Definition" section. The majority of the members commented that the COPM was "good" in incorporating clients' considerations of their own performance in all three areas of occupational performance. When asked to list the activities that were difficult to perform, clients tended to include activities from all these areas of self-care, productivity, and leisure.

Nevertheless, the major criticism of the content-related validity of the instrument is the uncertainty of the representativeness of the activities in reflecting a client's occupational performance dysfunction. Some of the panel members queried the accuracy and adequacy of using a semi-structured interview as the best procedure for acquiring clients' perceptions of their problems in performing the activities. The extent to which clients understood their own life situations, expectations, and functional status were thought to greatly influence their decisions. In addition, the panelists indicated that the therapists' interview techniques and philosophy of practice may also have had significant impact on the ways that clients responded. In general, the combination of these factors is likely to create dramatic variations in the choices of activities among the clients assessed by different clinicians using the COPM.

Suggestions for Change

1) Standardization of the testing process of the "Problem Definition" section would most likely reduce the variations of information gathered by different therapists during the interview. This could be achieved by introducing more guiding questions which therapists could use to prompt clients to review different areas of performance in self-care, productivity, and leisure. The authors of the COPM have suggested that a more structured interview violates the purposes and concepts of the client-centred model (Law et al., 1991). The authors argue that as long as the guiding questions

are explorative rather than restrictive and suggestive in nature, they will help therapists ensure that the opportunity for thorough exploration has been given to clients during the interview. However, some of the panel members in this study suggested showing a list of activities to clients which covered self-care, productivity, and leisure after the client had identified his/her activities. With this list, clients could then be requested to compare their own list to see if anything is missing in the activities that they have identified. In the case of clients who expressed difficulty in understanding words, written words could be replaced with pictures. These activities would be shown to the clients only after they had made their initial responses to the "Problem Definition" section of the COPM. The list of activities could be constructed by using Appendix C of the Test Manual. Extra activities could be added as was done in this study (increased to 108 activities).

2. Therapists who plan to administer the COPM should attend a workshop to learn the skills of semi-structured interviewing and to learn the use of the COPM (in addition to the seminar and training video that are available at the present time). More training in this area can improve the extent to which the testing process is standardized. An assessment of inter-rater reliability or agreement should also be incorporated into the training session. Videotaping the ways that different therapists administer the

COPM and evaluating their performance are possible ways to promote standardization across raters. The format and content of the training workshop provided by the FIM Group, Uniform Data System, Center for Functional Assessment Research in Buffalo (New York, United States) is a good example of the way that clinical observations and testing procedures can be standardized.

3) Before clients are asked to respond to the "Problem Definition" section of the COPM, therapists should provide a standardized information package for clients to complete. The package could include the following: medical information, prognosis of clients' impairment (if any), potential disability, and resources that are available to clients. Information of this kind is argued to be essential in building up clients' knowledge and insights into their own disabilities and potentials. Currently, educating and informing the clients depends on the preferences of the individual therapists and, therefore, the amount of "entry-knowledge" varies from client to client. More information provided to clients before conducting the COPM would minimize the undesirable effects of the COPM by giving clients a better understanding of their disabilities and environment.

COPM as a Measure of Performance Component

Results of this study revealed that the COPM was "fair" in measuring the performance components of clients. Comments from the panel members suggested that the instrument was not adequate in reflecting a client's dysfunction in the physical, mental, sociocultural, and spiritual components of the model of occupational performance. Despite the pressure in the profession to promote the assessment and treatment of occupational performance as the domain of concern (Fisher, 1992; Law, 1993; Trombly, 1993), any measurement of occupational performance should also reflect a client's dysfunction in the four performance components when, theoretically, the two levels of human performance have an interactive relationship (Christiansen, 1991; McColl & Pranger, 1994).

The measurement of both occupational performance and performance on the four performance components provides a clearer picture of a client's performance and the underlying causes contributing to his/her dysfunction. The authors of the COPM have recognized the importance of gathering both types of information before clinical reasoning can be made meaningful by fellow therapists. In the Test Manual, therapists are encouraged to "further assess performance components and the environment" with the use of other clinical instruments (Law et al., 1991, p.21) after clients have completed the COPM. However, it is argued that since the COPM is built on the occupational performance model, it is clinically non-meaningful if the results of assessments of the two levels are reviewed in isolation. There are many clinical enquiries which can not be answered in the existing COPM. For example, when a client is known to be unable to transfer into the bath tub, therapists want to know: why does it happen? and how can the problem be solved? A logical solution to this is to incorporate physical, mental, sociocultural, and spiritual assessment components into an extended version of the COPM.

Suggestions for Change

1. It is worthwhile to expand the COPM to incorporate sections that measure performance components. One way to do this is to construct new assessment items and sections that are relevant and representative of physical, mental, psychosocial, and spiritual constructs. Plenty of information is available in these areas, e.g. Christiansen and Baum's (1991) edited book Occupational therapy. Overcoming human performance deficits. However, this will involve a prolonged process and probably redundant process of instrumentation and validation since a lot of clinical instruments measuring different performance components are already available. Another method to expand the COPM is to adopt and include some of the well-established assessment batteries as part of the COPM assessment protocol. However, in-depth correlational and validation studies of the COPM with these other instruments would be needed before a well-grounded testing package of occupational performance would be available for clinical use.
2. More studies in the areas of occupational performance (CAOT, 1991), occupational science (Yerxa 1993; Yerxa et al., 1989), and human performance should be launched to build up the breadth and depth of the

theoretical models used in occupational therapy. Results of this study indicated the relationships between activity patterns and life roles of an individual. Data also revealed the close links between a person's self-perceived performance and his/her satisfaction with the performance. As shown in this study, validation of the COPM involved testing the model of occupational performance which ultimately contributed to the development of the theoretical framework. Better understanding of the theoretical constructs of occupational performance would enhance further improvement of the COPM and the interpretations of its scores.

COPM Incorporates Roles and Role Expectations of Clients

The results of this study have shown positive evidence that clients' responses made on the COPM reflect their existing life roles and role expectations. Activities selected by the clients in the "Problem Definition" section were found to match with their life roles as assessed by the Role Checklist. Protocol analyses of clients' responses gathered during the reassessment occasion suggested that clients used a variety of strategies when asked to identify the activities for the Performance and Satisfaction ratings in the COPM.

Suggestions for Change

1. It would be worthwhile to incorporate a new section before the "Problem Definition" section. The new section would ask clients to identify their

past life roles before their disabilities and future life roles after completion of rehabilitation program. This section would serve as a "warm-up" for clients to arouse their awareness of their life roles and role expectations before the "Problem Definition" section. Results of this new section could also be used to triangulate with clients' responses made in the "Problem Definition" section in order to ascertain the completeness of the "present" problems identified by the clients. The format of the assessment of life roles could be similar to that of the Role Checklist (Oakley et al., 1986) and would take only a short time to complete.

2. It is recommended that the "Problem Definition" process of the COPM be conducted every time the test is used to reassess clients rather than only at the initial assessment as is presently the case. Many clinicians have pointed out that, during the reassessment of clients, it is common to find clients reporting other activities which they perceived as difficult to perform, but which had not been identified in the initial assessment. In fact, this is a natural process as clients progress through occupational therapy intervention in which they experience "growth" and "change" when engaging in purposeful activities (CAOT, 1991; McColl & Pranger, 1994). New activities identified in the reassessment would be added to the original list of activities and be rated as to their Importance, perceived Performance, and Satisfaction with Performance.

COPM Incorporates Importance of Activities of Clients

A client-centred model recognizes the essence of treating clients as individuals who possess distinctive characteristics and needs. The consideration of clients' perceptions of what are and are not important in their lives is a major step forward in the measurement of clinical outcomes. The COPM has incorporated clients' perceptions by asking them to rate the activities identified in the "Problem Definition" section in terms of their importance in their lives. The Importance ratings act as weighting factors throughout the remainder of the scoring process (Law et al., 1991).

Results have illustrated the substantive validity of the Importance ratings. Empirical findings showed that activities that were regarded as more important by the clients were consistent with their life roles. The patterns of those activities were found to be similar to those patterns revealed by Law et al. (1994). The protocol analyses also revealed that when the clients were asked to rate the Importance of the identified activities, they managed to differentiate important activities from unimportant activities.

The two major drawbacks that have been revealed in this study with respect to "importance" in the COPM are: 1) the inaccuracy of using the ten-point rating scale (1 to 10) to differentiate activities which fall in the middle range of the scale, and 2) the methodological flaw of using the existing Importance ratings as weighting factors. Findings of this study suggested that the clients had difficulties in defining and differentiating activities which were mid-range or "less important", e.g. ratings of 7 or 8. Beyond ratings such as 10 (most important) or 4 (much less important), other ratings seemed to mean different things to different clients. The method of selecting activity items for the Performance and Satisfaction ratings is also challenged. According to the existing testing procedures, only the five activities having the highest Importance ratings are used for further ratings. This method greatly impedes the principle of weight assignment. First of all, it limits the variance of the Importance ratings to be used as weights. In this study, the mean Importance ratings of activities in different areas of performance ranged from 7.6 (SD = 2.3) to 9.8 (SD = 0.7). It is obvious that the restricted range of item weights likely did not have much impact on the final subscale scores on the COPM. In a recent pilot study, Law et al. (1994, p.196) stated:

> The COPM research team determined that scores on the COPM are equivalent whether or not importance weights are incorporated into the scoring and so these have been eliminated from the scoring formula.

However, it is argued that their "equivalence" was attributable to the restricted range imposed by the COPM testing procedures rather than by its psychometric property. In fact, according to Nunnally (1978), item weightings are useful if the number of items is less than 20. Furthermore, the incorporation of Importance weightings is one of the major principles to operationalize client-centred and occupational performance models.

Suggestions for Change

1. It is recommended that further investigations of the methods of measuring the construct of "Importance of activities" be undertaken. According to this study, clients' decisions on the ratings of importance of activities were much more complicated than anticipated. Besides the ten-point rating scale, other scaling methods such as the semantic-differential scale (Nunnally, 1978) may better reflect the perception and values of clients.
2. Another method to improve the usefulness of the Importance ratings would be to incorporate the Importance ratings made by both rehabilitation professional and "significant-others" of the clients. The clinical experience of therapists has revealed that clients' decisions may be valid with respect to their life situation before their disabilities, but not their present or future situations. Composite ratings of clients, therapists, and significant-others could further enhance the meaningfulness of the existing weighting system.
3. In order to accommodate changes in clients' perceptions of Importance of particular activities throughout the process of rehabilitation, the Importance ratings, as with the "Problem Definition", should be revised every time the clients are assessed with the COPM. Discrepancies revealed between the assessments at different stages of treatment provide valuable information about the extent to which clients perceive changes in their life roles and

expectations, and their environments as they progress through occupational therapy intervention.

4. More than five activities should be included in the ratings of perceived performance and satisfaction with that performance. In the Test Manual, no information was given on how the number of five was determined. Based on the findings of this study, the average number of activities identified by the clients was 8. In Law et al.'s (1994) study, 49.6% of the clients identified 5 or more activities. It was found that the limit of five activities greatly restricted the range of the weighting system making it non-discriminating. Furthermore, a lot of information was missed on clients' occupational performance with a limit of five activities. As a consequence, this restriction continues to pose a major threat to the content-related validity of the interpretation of the COPM results.

COPM Measures Client's Self-Perceived Performance

Findings suggest that the construct of Performance has been well demonstrated in the COPM. The protocol analysis of clients' responses showed that most of the clients understood the concept of performance. However, the accuracy and meaningfulness of assigning Performance ratings using the 10-point scale were challenged. The evidence of convergent validity between the Performance subscale scores and the KBADL and FIM was weak.

Suggestions for Change

1. The ten-point rating scale may not be the most relevant method in reflecting clients' performance. The reason for choosing this scale was not well explained in The Manual. As a matter of fact, studies on the validity of self-report assessment have not revealed the superiority of using this particular type of scale. An alternative may be to use an ordinal rating scale with consecutive levels of clients' competence. One such scale is the rating scale used in the Functional Independence Measure. Instead of ten points, there are seven points on the scale scored from one to seven. As applied to the COPM, the two anchor points would remain the same as "not able to do it" and "able to do it well". However, the five ratings in the middle would be differentiated by the amount of assistance needed from others with ranges including "modified independence with device" to "maximal assistance".

2. Before clients assign the Performance ratings, opportunities should be provided for them to perform these activities with guidance from therapists. Many of the clients who participated in the study commented that they did not know how to evaluate their own performance without actually "doing" the activities. In the present testing protocol, for those clients who have not had a chance to try out the activities, they are required to imagine their performance in those activities. This problem was found to be especially

serious in the initial assessment where most of the clients had to rely on their own subjectivity and imagination.

3. Although clients' self-report ratings were rated as useful and accurate in the literature (Elam et al., 1991, Myer et al., 1993), a better clinical inference could be achieved if the COPM incorporated ratings of clients' performance by therapists on the activities items. Similarities between clients' self-rating and those by therapists would confirm the accuracy of clients' rating. On the other hand, discrepancies between the two sets of ratings would serve to arouse communication between therapists and clients on the reasons for those discrepancies. Some of these reasons could be: differences in expectations, differences in definitions of competence and performance, and effects of motivation.
4. As in the Importance rating, more activities should be included for Performance ratings to minimize the threats to content-related validity of the COPM results.

COPM Measures Client's Satisfaction with Performance

Incorporation of clients' satisfaction in the quantification of occupational performance is consistent with the theoretical construct that suggests that fulfilment and satisfaction are the outcomes when individuals' expectations are met. The protocol analysis revealed that the clients who participated in this study used different strategies in rating their satisfaction. These strategies were compatible with those suggested in existing theories describing the phenomenon. However, the major problem with the Satisfaction items and subscale scores was the inaccurate ratings that resulted from the use of the ten-point rating scale. Convergent validity established between clients' pre-discharge COPM Satisfaction subscale scores and those on the SPSQ ranged from -.13 to .36.

Suggestions for Change

1. More exploration in the structure of the construct of Satisfaction is required. The notion that satisfaction is a unidimensional construct assumed by the COPM is not well justified. Besides the use of a ten-point rating scale, other scaling methods such as a semantic-differential scale should be considered.
2. Like the Performance subscale, the number of activities to be rated should be increased in order to improve reliability and minimize the threats to content-related validity of the COPM assessment results.

Use of Ten-Point Rating Scale

In the COPM, clients' Importance, Performance, and Satisfaction are rated with the same type of scale. Clients are asked to rate their own values and evaluations using ten-point rating scales. Each rating scale has two end-points which represent two opposite anchor points of the constructs of importance, performance, and satisfaction. The two anchor points are assigned 1 and 10 which represent an increase in quantity of the construct as the scale progresses from the smaller to the larger number. Between 1 and 10, the scale is further divided into eight equal portions with numbers from 2 to 9 assigned to each point. No qualifiers nor criteria are associated with the middle eight numbers.

The ten-point scales used in the COPM assume unidimensional ordinal constructs of the attributes rated by clients. As argued earlier in this chapter, this assumption has not been substantiated, especially in the area of satisfaction where unidimensionality may over-simplify the structural fidelity of the construct measured. The disadvantages of assigning numbers without qualifiers on an ordinal scale are obvious. As Nunnally (1978) commented, the concepts of number may vary between individuals. Besides the two anchor points, numbers of 2 to 9 may be interpreted differently from client to client. The memory carried over from the initial assessment is another common undesirable effect in using rating scales of this kind.

Suggestions for Change

1. The existing ten-point rating scales should be replaced with visual analogue scales which do not have numerical labels. Clients would be asked to assign a point on the visual analogue scale between the same two existing anchor points. The marked point would then be measured against a particular metric scale, e.g. in centimeters or inches, and the actual distances reported. Ratings on visual analogue scales are regarded as being at a continuous level of measurement (DeVellis, 1991).

2. If the existing ordinal scale is used, the number of segments in the scales should be reduced to five or seven. In fact, it has been suggested that the gain in reliability of the scale by increasing the number of scale steps levels off at about seven (Nunnally, 1978). Moreover, the more the steps in the scale, the more difficulty clients have in deciding on their ratings. In order to improve the reliability and truthfulness of the scale, it is recommended that a few scenarios be constructed for clients to practise using the scales before they use them for the actual ratings. This pre-rating practice would help clients locate the two anchor points of the scales with their subjective experience on Importance, Performance, and Satisfaction. This method is identical to various types of calibration processes adopted by medical and physical clinical instruments.
3. Since the existing scales of the COPM do not possess analogue characteristics, it would be useful to develop descriptors and criteria for each numbered point on the scale. This would help further standardize client's ratings and make inter-individual comparisons with the COPM scores more meaningful.

Scoring System of COPM

The COPM yields items and subscale scores for the Performance and Satisfaction subscales. These two subscale scores are obtained for both initial and re-assessment, and are compared by computing the Change in Performance and

Satisfaction scores. Therapists are advised to interpret those "Change" scores which reflect changes of clients' occupational performance attributable to occupational therapy interventions.

In computational terms, the Importance ratings were not found to exert any effects when the Performance and Satisfaction subscale scores were compared intra-individually since the same set of weights is applied to the same activity items to yield the two subscale scores. Hence, intra-individual comparisons such as the Change in Performance and Satisfaction scores, and the subscale scores can be conducted without considering the Importance score. In fact, Law et al.(1994) has recommended eliminating the Importance rating from the scoring formula altogether. However, the underlying reason for the claim was not substantiated in the article.

If the recommendations on repeating the "Problem Definition" and Importance ratings in every re-assessment are adopted, the incorporation of Importance ratings in the computation of subscale scores would be meaningful. It is further argued that individual clients' Importance ratings would be important if the COPM subscale scores were used for inter-individual comparisons. In norm-referenced assessment, the Importance ratings of clients assigned to different activities are reflective of the individual client's role expectations and environmental characteristics. Importance ratings, therefore, would become essential if comparisons were conducted among different clients and different diagnostic groups. Although the Test Manual explicitly specifies that the COPM is an individualized criterion-referenced assessment, it is useful to explore the feasibility of using the COPM scores as a norm-referenced measure if it is intended to be used as an outcome measure for evaluating efficacy of clinical interventions.

The computation of Performance and Satisfaction subscale scores from their respective item scores was also challenged. Empirical findings revealed very high correlations between the item Performance and Satisfaction scores (ranging from .37 to .98). This is consistent with those reported by Law et al. (1994) which was .76 ($p < .001$). Moreover, the stability estimates of the Performance and Satisfaction subscale scores were as low as .32 and .09 respectively. This further demonstrated that the COPM subscale scores were unstable under different assessment occasions and different testers. These errors greatly impede the meaningfulness of the subscale scores on the COPM, and in turn, their interpretations. Since the COPM's subscale scores are problematic, their derivatives of the Change of Performance and Satisfaction scores are also expected to be problematic.

Suggestions for Change

1. If the Importance ratings are to continue to be incorporated in the COPM, the present strategy of using the ratings should be altered. In the existing testing protocol, the Importance ratings are taken as absolute values. This is problematic if comparisons are made between different clients who select

different numbers of activities. Changing clients' Importance ratings from absolute values to relative values would resolve part of the problem. Relative values are assigned to particular activities depending not only on the Importance ratings of specific activity items but also the total value of Importance ratings made to all the activities items.

Table 7.1

Relative Importance Index of the COPM

Activities Identified	Importance Ratings	Relative Importance Index
Dressing Upper Garment	8	0.24
Toiletting	8	0.24
Cooking a Meal	7	0.21
Car Transfer	4	0.12
Gardening	6	0.18
Total Importance Rating	33	1.00

Table 7.1 is an example of the Importance ratings transformed to a "relative Importance index" of a client (by dividing each Importance rating by the total Importance rating). The relative Importance index does not

have an effect on intra-individual comparisons, such as between initial and pre-discharge subscale scores because it is simply a rescaling of the original Importance ratings. However, it does affect the values of the subscale scores when two individual clients, or two groups of clients, are compared.

Table 7.2

Item Performance Scores Using Importance Rating and Relative Importance Index

Activity Items	Import. Rating		Relative Importance Index		Perf. Rating		Item Perf. Score		Adjusted Item Perf. Score	
Clients	A	B	A	B	A	B	A	B	A	B
Dressing	8	6	0.24	0.22	5	5	40	30	1.20	1.10
Toiletting	8	6	0.24	0.22	4	4	32	24	0.96	0.88
Cooking	7	5	0.21	0.19	4	4	28	20	0.84	0.76
Car Transfer	4	3	0.12	0.11	2	2	8	6	0.24	0.22
Gardening	6	7	0.18	0.26	2	2	12	14	0.36	0.52
Total	33	27	1.00	1.00			120	94	3.60	3.48

In Table 7.2, the Performance subscale and adjusted item Performance subscale scores of clients A and B are compared. The adjusted item Performance score is the product of multiplying perceived Performance ratings by the relative Importance index. In the table, clients A and B have similar if not identical ratings of their perceived performance on the five activity items (refer to the Perf. Rating column).

In the adjusted item Performance scores column, scores are much smaller than the respective Performance ratings due to its rescaling. The Performance subscale scores of client A is 24 (120 divided by the number of items which is 5), whereas client B is 18.8 (94 divided by 5), with a difference of 21.6%. The adjusted Performance subscale score of client A is 0.72 (3.60 divided by 5), whereas client B is 0.70 (3.48 divided by 5), with a difference of only 2.8%. When the adjusted item Performance scores are used, the differences in the Performance subscale scores between the two clients become much smaller and reflect less influence from client A's use of larger values in the Importance ratings. The difference in using the original Importance rating and the relative Importance index hence are evident.

The use of the relative Importance index accommodates the differences in numbers of activities clients identify due to variances in their performance components including capabilities, life roles and expectations, and environmental demands. The relative Importance index reflects the different emphases that individual clients put on different activities which is an important component of occupational performance. This same process can also be applied to the adjustment of item Satisfaction scores and hence producing adjusted Satisfaction subscale scores similar to the adjusted Performance subscale scores. An improved method of using the Importance ratings has been offered, but the variations in clients' performance in different activities still does not allow meaningful comparisons of occupational performance between individual clients or client groups. This will be discussed under point 2 below.

2. The evaluation of human performance in occupational therapy is problematic because of its diversity and complexity. Inter-individual comparisons are difficult to conduct because of the variations in abilities and activities identified by each individual. In clinical situations, the basis upon which clients are compared is controversial. In order to make clinical assessment simple, various instruments have been built on the "minimum item approach". These instruments, such as the Functional Independence Measure (Granger, Hamilton, & Sherwin, 1988) and Barthel Index (Mahoney & Barthel, 1965), have dominated the field for some time. The use of these instruments assumes that the tasks assessed are essential to clients and that they are good predictors of a client's independence. Many

commentators have pointed out problems with the content-related validity of these instruments when minimum items are used to assess clients' performance results (Christiansen, 1993; Law, 1993; Trombly, 1993).

In response to these criticisms, the applications of item response theory (Hambleton, 1993; Hambleton, Swaminathan, & Rogers, 1991) and many-faceted Rasch analysis (Dickerson & Fisher, 1993; Fisher, 1993; Linacre, 1988, 1989) in functional performance evaluation have become popular in the recent few years. The basic principles underlying the applications are two fold. First, individuals' performance can be accounted for by a set of traits which are the individuals' abilities; and second, the relationships between clients' performance and their abilities can be modelled by different mathematical functions. The advantage of using item response theory in functional performance evaluation is that it allows calibration of different activities (items) performed by different clients on the same scale of clients' abilities (traits).

Different activities can be assigned with specific values of "item difficulty", i.e. position on the independence continuum. In addition, clients' performance in different activities allows the estimation of their abilities (level of independence). With the application of item response theory, individual clients' performance can be compared on the basis of the different activities performed.

The application of item response theory has shed some light on the way in which the COPM could be adapted to allow for inter-individual comparisons and norm-referenced testing. In the COPM, individual clients are expected to identify different activities based on their different abilities, life expectations, and environmental demands. This has been one of the major obstacles for making inter-individual comparisons. To begin the process, activities which are commonly selected by clients would be calibrated in terms of clients' abilities by the many-faceted Rasch analysis. Since clients with different disabilities may perform the same activity differently, it is important to keep the sample with which activities are calibrated as homogeneous as possible (e.g. clients with lower limb orthopedic problems, or clients with arthritis involving the upper limb joints). Similar to the present COPM testing procedures, clients are asked to assign the Importance ratings and rate their own performance in the selected activities. Different from the present COPM procedures, therapists ratings on clients' performance would also be required. The final item Performance ratings are then reached through constructive discussion between the client, therapist, and significant-others of clients. Item Performance scores, incorporating item weights in terms of a relative Importance Index, are entered to Rasch-based computer program for the calibration of item difficulties. Subsequently, occupational performance of clients in terms of logit (ability) scores would be estimated.

There are many issues that need to be resolved before a norm-referenced COPM can be realized. First of all, the dimensionality and additivity of Performance ratings needs to be affirmed. Second, the mechanism with which the relative Importance index can be incorporated in the difficulty levels of the items has not been investigated. Third, the actual operations that are involved in obtaining the estimates of clients' abilities by Rasch analysis must be explored in great detail.

3. It is recommended that the Satisfaction subscale be removed from the COPM assessment protocol. This is supported by the observations that the Satisfaction and Performance subscales are highly correlated. From a measurement perspective, this means that the use of the Satisfaction subscale does not seem to increase the amount of information on the evaluation of clients' occupational performance. As a matter of fact, analysis of clients' protocols suggested that their satisfaction is closely related to their perceived performance. However, since satisfaction is an important component of occupational performance, it is suggested that a global satisfaction scale be constructed to obtain clients' feedback on their performance using a visual analogue scale for the different areas of occupational performance, i.e. self-care, productivity, and leisure.

4. The Change in Performance and Satisfaction scores should no longer be used without strong evidence on acceptable error of measurement and reliability estimates. The application of Rasch analysis manages to bridge the gap between intra- and inter-individual comparisons of clients' occupational performance. Ultimately, the inter-individual comparisons could form the basis for measuring the progress of clients and setting discharge criteria, whilst the intra-individual comparisons could form the basis for both formative and summative evaluation of clinical programs.

Reliability of COPM

The plan for estimating the reliability of the COPM turned out to be ill-concerned in the light of the profound changes that occured in the clients over a short period of time. As mentioned earlier, responses of the clients on the COPM varied dramatically during the three day (for orthopedic) to seven day (for stroke) retest periods. These variations confounded by the treatment and tester effects were impossible to be disentangled.

Future research should be pursued in developing valid methods in establishing the reliability of the COPM. An accurate estimate of inter-tester reliability could be achieved with generalizability studies (Cronbach, Gleser, Nanda, & Rajaratnam, 1972). With Rasch analysis, the dimensionality, scalability, and internal consistency of the COPM could also be obtained. The appropriateness of documenting test-retest reliability as evidence of psychometric properties of clinical instrument requires further investigations.

Universal Utility of COPM

The data gathered in this study do not allow for a thorough exploration of the extent to which the COPM can be used across the spectrum of clients with different disabilities. The small sample sizes of the orthopedic and stroke groups, especially the latter, failed to provide generalized results on how much the responses made by the clients in the two groups differed. Initial quantitative and qualitative data analyses suggested that significant discrepancies in the responses of the two groups did not occur. More research is called for to investigate the utility of COPM in clients with different disabilities. However, it is argued that the use of the COPM with proxies and with those who have either mental or cognitive deficits is likely to pose major threats to the validity of the COPM assessment results. More research is required before the utility of the COPM can be widened to all client groups.

Conclusion

The Canadian Occupational Performance is built on the Canadian model of occupational performance. Since its publication in 1991, it has aroused interests in exploring its psychometric properties and clinical utilities in occupational therapy. Evidence gathered in this study has shown both the positive and negative perspectives of using the COPM to measure occupational performance of clients in the orthopedic and stroke groups. Empirical findings have shown that the COPM incorporates the model of occupational performance, especially in its ratings of importance, performance, and satisfaction. At a conceptual level, the COPM has been successful in its reflection of occupational performance. The notion that occupational performance is the product of the interactions of individuals' performance components and their environment has found support. However, at a clinical level, the instrument has been unable to provide adequate information to directly enhance meaningful clinical reasoning of therapists. This is attributed to the weaknesses in the test construct, the test content, the test dimensionality, and the testing processes of the COPM with which the scores reflecting clients' occupational performance are yielded. Problems in its instrumentation have led to difficulties in interpreting the COPM scores, and hence clinicians' inferences made about clients' performance tend to be inaccurate.

The assessment format of using the face-to-face interview is a strength of the COPM. However, more stringent training and standardization of methods are needed so that more consistent results can be obtained to ensure better test-retest and inter-tester reliability of the COPM scores.

Nonetheless, the publication of the Canadian Occupational Performance Measure is a major step forward in operationalizing client-centred and occupational performance-based assessment in occupational therapy. Evidence of construct validity gathered in this study is regarded as a means rather than an end to the betterment of the instrument in measuring clients' occupational performance. Further research and effort should be directed toward improving the instrument. A better use and interpretation of the assessment results can contribute to the day-to-day practice of occupational therapist, the profession, and finally, to clients as consumers of occupational therapy services.

REFERENCES

Alberta Association of Registered Occupational Therapists. (1990). Standard of practice. Edmonton, AB: AAROT.

American Psychological Association. (1954). Technical recommendations for psychological tests and diagnostic techniques. Psychological Bulletin, 51, 201-238.

American Psychological Association, American Educational Research Association, & National Council on Measurement in Education. (1985). Standards for educational and psychological testing. Washington, DC: American Psychological Association.

Anastasi, A. (1982). Psychological testing (5th ed.). New York: MacMillan Publishing Co.

Anderson, J. R., & Bower, G. H. (1973). Human associative memory. Washington, DC: Winston.

Benson, J., & Clark, F. (1982). A guide for instrument development and validation. American Journal of Occupational Therapy, 36(12), 789-800.

Betz, N. E., & Weiss, D. J. (1976). Validity. In B. Bolton (Ed.)., Handbook of measurement and evaluation in rehabilitation (pp.39-60). Baltimore: University Park Press.

Cammermeyer, M., & Evans, J. (1988). A brief neurobehavioral exam useful for early detection of postoperative complications in neurosurgical patients. Journal of Neuroscience Nursing, 20(5), 314-323.

Campbell, D. T., & Fiske, D. W. (1959). Convergent and discrimination validation by the multitrait-multimethod matrix. Psychological Bulletin, 56, 81-105.

Campbell, D. T., & Fiske, D. W. (1967). Convergent and discrimination validation by the multitrait-multimethod matrix. In D.N. Jackson and S. Messick (Eds.), Problems in human assessment (Chapter 6, pp.124-132). New York: McGraw Hill Book Co.

Campbell, S. K. (1989). Measurement in developmental therapy: Past, present, and future. Physical and Occupational Therapy in Pediatrics, 9(1), 1-14.

Canadian Association of Occupational Therapists. (1991). Occupational therapy guidelines for client-centred practice. Toronto, ON: CAOT Publications.

Carp, F. M., & Carp, A. (1982). Test of a model of domain satisfaction and well being: Equity considerations. Research on Aging, 4, 503-522.

Carswell, A., Carson, L. J., Walop, W., & Zgola, J. (1992). A theoretical model of functional performance in persons with Alzheimer disease. Canadian Journal of Occupational Therapy, 59(3), 132-140.

Center for Functional Assessment Research. (1991). Guide for the use of the Uniform Data Set for Medical Rehabilitation. Version 3.1. Buffalo: Research Foundation - State University of New York.

Chamberlain, K. (1988). On the structure of well being. Social Indictors Research, 20, 518-604.

Christiansen, C. (1991a). Occupational therapy. Intervention of life performance. In C. Christiansen and C. Baum (Eds.), Occupational therapy. Overcoming human performance deficits (Chapter 1, pp. 1-43). Thorofare, NJ: SLACK Inc.

Christiansen, C. (1991b). Occupational performance assessment. In C. Christiansen and C. Baum (Eds.), Occupational therapy. Overcoming human performance deficits (Chapter 15, pp.375-426). Thorofare, NJ: SLACK Inc.

Christiansen, C. (1993). Continuing challenges of functional assessment in rehabilitation: Recommended changes. American Journal of Occupational Therapy, 47(3), 258-259.

Christiansen, C., & Baum, C. (Eds.). (1991). Occupational therapy. Overcoming human performance deficits. Thorofare, NJ: SLACK Inc.

Christiansen, C., Schwartz, R., & Barnes, K. (1988). Self-care: Evaluation and management. In DeLisa, J. (Ed.), Rehabilitation medicine: Principles and practice (pp. 95-115). Philadelphia: J.B. Lippincott.

Clark, P. (1979a). Human development through occupation: Theoretical framework in contemporary occupational therapy practice. Part 1. American Journal of Occupational Therapy, 33, 505-514.

Clark, P. (1979b). Human development through occupation: Theoretical framework in contemporary occupational therapy practice. Part 2. American Journal of Occupational Therapy, 33, 577-585.

Cohen, J. A. (1960). A coefficient of agreement for nominal scales. Educational and Psychological Measurement, 20, 37-46.

Cohen, J. (1977). Statistical power analysis for the behavioral sciences (rev.ed.). New York: Academic.

Cook, D. C., & Campbell, D. T. (1979). Quasi-experimentation. Design & analysis issues for field settings. Chicago: Rand McNally College Publishing Company.

Crocker, L., & Algina, J. (1986). Introduction to classical and modern test theory. Orlando, Florida: Holt, Rinehart and Winston, Inc.

Cronbach, L. J. (1971). Test validation. In R.L. Thorndike (Ed.), Educational Measurement (2nd ed.). (Chapter 14, pp. 443-507). Washington, DC: American Council on Education.

Cronbach, L. J. (1984). Essentials of psychological testing (4th ed.). New York: Harper & Row, Publishers.

Cronbach, L. J. (1988). Five perspectives on validation argument. In H. Wainer & H. Braun (Eds.), Test validity (pp.3-17). Hillsdale, NJ: Erlbaum.

Cronbach, L. J., Gleser, G. C., Nanda, H., & Rajaratnam, N. (1972). The dependability of behavioral measurements. New York: John Wiley.

Cronbach, L. J., & Meehl, P. E. (1955). Construct validity in psychological tests. Psychological Bulletin, 52, 281-302.

Davidson, H. (1991). Performance and the social environment. In C. Christiansen and C. Baum (Eds.), Occupational therapy. Overcoming human performance deficits (Chapter 6, pp.144-177). Thorofare, NJ: SLACK Inc.

Department of National Health and Welfare and Canadian Association of Occupational Therapists. (1983). Guidelines for the client-centred practice of occupational therapy (H39-33/1983E). Ottawa, ON: Department of National Health and Welfare.

Department of National Health and Welfare and Canadian Association of Occupational Therapists. (1986). Intervention guidelines for the client-centred practice of occupational therapy (H39-100/1986E). Ottawa, ON: Department of National Health and Welfare.

Deusen, J. V., Shalik, L., & Harlowe, D. (1990). Construct validation of an acute occupational therapy cerebral vascular accident assessment tool. Canadian Journal of Occupational Therapy, 57(5), 155-159.

DeVellis, R. F. (1991). Scale development. Theory and application. Newbury Park, CA: SAGE Publications Inc.

Dickerson, A. E., & Fisher, A. G. (1993). Age differences in functional performance. American Journal of Occupational Therapy, 47(8), 686-692.

Dunn, W. W. (1989). Validity. Physical and Occupational Therapy in Pediatrics, 9(1), 149-168.

Dunn, W. (1993). Measurement of function: Action for the future. American Journal of Occupational Therapy, 47(4), 357-359.

Ebel, R. L. (1967). Evaluating content validity. In D.A. Payne & R.F. McMorris (Eds.), Educational and psychological measurement: Contributions to theory and practice (pp.85-94). Waltham, MA: Blaisdell.

Edwards, M. M. (1990). The reliability and validity of self-report activities of daily living scales. Canadian Journal of Occupational Therapy, 57(5), 273-278.

Elam, J. T., Graney, M. J., Beaver, T., Derwi, D. E., Applegate, W. B., & Miller, S. T. (1991). Comparison of subjective ratings of function with observed functional ability of frail older people. American Journal of Public Health, 81(9), 1127-1129.

Ericsson, K. A., & Simon, H. A. (1980). Verbal reports as data. Psychological Review, 87(3), 215-248.

Ericsson, K. A., & Simon, H. A. (1985). Protocol analysis. Handbook of Discourse Analysis. Vol. 2 (Chapter 14, pp.259-268). London: Academic Press.

Evan, J., & Salim, A. A. (1992). A cross-cultural test of the validity of occupational therapy assessments with patients with schizophrenia. American Journal of Occupational Therapy, 46(8), 685-695.

Exner, C. E. (1993). Content validity of the in-hand manipulation test. American Journal of Occupational Therapy, 47(6), 505-513.

Fillenbaum, G. G., & Smyer, M. A. (1981). The development, validity, and reliability of the OARS Multidimensional Functional Assessment Questionnaire. Journal of Gerontology, 36(4), 428-434.

Fisher, A. G. (1992). Functional measure, Part I: What is function, what should be measure, and how should we measure it? American Journal of Occupational Therapy, 46(2), 183-185.

Fisher, A. G. (1993). The assessment of IADL motor skills: An application of many-faceted Rasch analysis. American Journal of Occupational Therapy, 47, 319-329.

Fisher, G. H., & Short-DeGraff, M. (1993). Improving functional assessment in occupational therapy: Recommendations and philosophy for change. American Journal of Occupational Therapy, 47(3), 199-201.

Fricke, J., Unsworth, C., & Worrell, D. (1993). Reliability of the Functional Independence Measure with occupational therapists. Australian Occupational Therapy Journal, 40(1), 7-15.

Fuhrer, M. J., Rintala, D. H., Hart, K, A., Clearman, R., & Young, M. E. (1992). Relationship of life satisfaction to impairment, disability, and handicap among persons with spinal cord injury living in the community. Archieves of Physical Medicine and Rehabilitation, 73, 552-557.

Giles, G. M. (1994). Illness behavior after severe brain injury: Two case reports. American Journal of Occupational Therapy, 48(3), 247-250.

Glass, G. V., & Hopkins, K. D. (1984). Statistical methods in education and psychology (2nd ed). Massachusetts: Allyn and Bacon.

Granger, C. V., Hamilton, B. B., Linacre, J. M., Heinemann, A. W., & Wright, B. D. (1993). Performance profiles of the Functional Independence Measure. American Journal of Physical Medicine and Rehabilitation, 72, 84-89.

Granger, C. V., Hamilton, B., & Sherwin, F. (1988). Functional Independent Measure. Buffalo: State University of New York Press.

Guion, R. M. (1974). Open a window: Validities and values in psychological measurement. American Psychologist, 29, 287-296.

Gulliksen, H. (1950). Theory of mental tests. New York: Wiley.

Hallett, J. D., Zasler, N. D., Maurer, P., & Cash, S. (1994). Role changes after traumatic brain injury in adults. American Journal of Occupational Therapy, 48(3), 241-246.

Hambleton, R. K. (1993). Principles and selected applications of item response theory. In R.L. Linn (Ed.), Educational measurement (3rd. ed.) (Chapter 4, pp.147-200). Phoenix, AZ: American Council on Education and The Oryx Press.

Hambleton, R. K., Swaminathan, H., & Rogers, H. J. (1991). Fundamentals of item response theory. California: SAGE Publications.

Hamilton, B. B., Granger, C. V., Sherwin, F. S., Zielezny, M., & Tashman, J. S. (1987). A uniform national data system for medical rehabilitation. In Fuhrer, M.J. (Ed.), Rehabilitation outcomes. Analysis and measurement (pp.137-147). Baltimore: Paul.H. Brookes Publishing Co.

Heinemann, A.W., Linacre, J. M., Wright, B. D., Hamilton, B. B., & Granger, C. (1993). Relationships between impairment and physical disability as measured by the Functional Independence Measure. Archives of Physical Medicine and Rehabilitation, 74, 566-572.

Hopkins, H. L., & Smith, H. D. (1988). Willard and Spackman's occupational therapy (7th ed.). Philadelphia: J.B. Lippincott Co.

Houston, D., Williams, S. L., Bloomer, J., & Mann, W. C. (1989). The Bay Area Functional Performance Evaluation: Development and standardization. American Journal of Occupational Therapy, 43(3), 170-183.

Howe, M. C., & Briggs, A. K. (1982). Ecological systems model for occupational therapy. American Journal of Occupational Therapy, 36, 322-327.

Katz, N., Itzkovich, M., Averbuch, S., & Elazar, B. (1989). Loewenstein Occupational Therapy Cognitive Assessment (LOTCA) Battery for brain injured patients: reliability and validity. American Journal of Occupational Therapy, 43(3), 184-192.

Kielhofner, G. (1983). Health through occupation: Theory and practice in occupational therapy. Philadelphia: F.A. Davis Co.

Kielhofner, G. (1985). A model of human occupation. Baltimore: Williams and Wilkins.

Kielhofner, G. (1992). Conceptual foundations of occupational therapy. Philadelphia: F.A. Davis.

Kielhofner, G., & Burke, J. P. (1980). Components and determinants of human occupation (pp.12-36). In Kielhofner, G.W.(Ed.), A model of human occupation: Theory and application. Baltimore: Williams and Wilkins.

Kiernan, R. J., Meuller, J., Langston, J. W., & Van Dyke, C. (1987). The Neurobehavioral Cognitive Status Examination: A brief differentiated approach to cognitive assessment. Annals to Internal Medicine, 107, 481-485.

Klein, R. M., & Bell, B. (1979). The Klein-Bell ADL Scale Manual. Seattle, Washington; University of Washington Medical School, Health Science Resource Center/SB-56.

Klein, R. M., & Bell, B. (1982). Self-care skills: Behavioral measurement with the Klein-Bell ADL Scale. Archives of Physical Medicine and Rehabilitation, 63, 335-338.

Krefting, L. H., & Krefting, D. V. (1991). Culture influences on performance. In C. Christiansen and C. Baum (Eds.), Occupational therapy. Overcoming human performance deficits (Chapter 4, pp.102-122). Thorofare, NJ: SLACK Inc.

Law, M. (1987). Measurement in occupational therapy: Scientific criteria for evaluation. Canadian Journal of Occupational Therapy, 54(3), 133-138.

Law, M. (1993). Evaluating activities of daily living: Directions for the future American Journal of Occupational Therapy, 47(3), 233-237.

Law, M., Baptiste, S., Carswell-Opzoomer, A., McColl, M., Polatajko, H., & Pollock, N. (1991). Canadian Occupational Performance Measure Manual. Toronto, ON: CAOT Publications.

Law, M., Baptiste, S., McColl, M., Opzoomer, A., Polatajko, H., & Pollock, N. (1990). The Canadian occupational performance measure: An outcome measure for occupational therapy. Canadian Journal of Occupational Therapy, 57(2), 82-87.

Law, M., Polatajko, H., Pollock, N., McColl, M. A., Carswell, A., & Baptiste, S. (1994). Pilot testing of the Canadian Occupational Performance Measure: Clinical and measurement issues. Canadian Journal of Occupational Therapy, 61(4), 191-197.

Linacre, J. M. (1988). FACETS: Computer program for many-faceted Rasch measurement. Chicago: MESA.

Linacre, J. M. (1989). Many-faceted Rasch measurement. Chicago: MESA.

Llorens, L. A. (1991). Performance tasks and roles throughout the life span. In C. Christiansen and C. Baum (Eds.), Occupational therapy. Overcoming human performance deficits (Chapter 2, pp.46-66). Thorofare, NJ: SLACK Inc.

Loevinger, J. (1957). Objective tests as instruments of psychological theory. Psychological Reports, 3, 635-694 (Monograph Supp. 9).

Loevinger, J. (1967). Objective tests as instruments of psychological theory. In D.N. Jackson and S. Messick (Eds.), Problems in human assessment (Chapter 5, pp.78-123). New York: McGraw Hill Book Co.

Mahoney, F. I., & Barthel, D. W. (1965). Functional evaluation: The Barthel Index. Maryland State Medical Journal, 14, 61-65.

Maslow, A. H. (1970). Motivation and personality (2nd ed.). New York: Harper & Row.

Mathiowetz, V. (1993). Role of physical component evaluations in occupational therapy functional assessment. American Journal of Occupational Therapists, 47(3), 225-230.

McCall, R. B. (1975). Fundamental statistics for psychology. New York: Harcourt, Brace.

McColl, M. A., & Pranger, T. (1994). Theory and practice in the Occupational Therapy Guidelines for Client-Centred Practice. Canadian Journal of Occupational Therapy, 61(5), 250-259.

McDermott, P. A., & Watkins, M. W. (1979). A program to evaluate general and conditional agreement among categorical assignments of many raters. Behavior Research Methods and Instrumentation, 11(3), 399-400.

Meadow, H. L., Mentzer, J. T., Rahtz, D. R., & Sirgy, M. J. (1992). A life satisfaction measure based on judgement theory. Social Indicators Research, 26, 23-59.

Meehl, P. E., & Rosen, A. (1956). Antecedent probability and the efficiency of psychometric signs, patterns, or cutting scores. Psychological Bulletin, 52, 194-216.

Messick, S. (1980). Test validity and ethics of assessment. American Psychologist, 35, 1012-1027.

Messick, S. (1989). Validity. In R. L. Linn (Ed.), Educational Measurement (3rd. ed.) (pp.13-104). New York: American Council on Education and Macmillan.

Messick, S. (1993). Validity. In R. L. Linn (Ed.), Educational measurement (3rd ed.)(pp.13-103). Phoenix, AZ: Oryx Press.

Meyers, J., Lytle, S., Palladino, D., Devenpeck, G., & Green, M. (1990). Think-aloud protocol analysis: An investigation of reading comprehension strategies in fourth- and fifth-grade students. Journal of Psychoeducational Assessment, 8, 112-127.

Michalos, A. (1980). Satisfaction and happiness. Social Indictors Research, 8, 385-422.

Mosey, A. C. (1986). Psychosocial components of occupational therapy. New York: Raven Press.

Myers, A. M., Holliday, P. J., Harvey, K. A., & Hutchinson, K. S. (1993). Functional performance measures: Are they superior to self-assessments? Journal of Gerontology, 48(5), M196-206.

Neter, J., Wasserman, W., & Kutner, M. H. (1990). Applied linear statistical models(3rd ed.). Tokyo: Toppan Co.

Newell, A., & Simon, H. A. (1972). Human problem solving. Englewood Cliffs, NJ: Prentice- Hall.

Northern California Neurobehavioral Group. (1988). Manual for the Neurobehavioral Cognitive Status Examination. Fairfax, CA: The Northern California Neurobehavioral Group, Inc.

Nunnally, J. C. (1978). Psychometric Theory (2nd ed.). New York: McGraw-Hill.

Oakley, F. (1982). The model of human occupation in psychiatry. Unpublished master's thesis, Virginia Commonwealth University, Richmond, Virginia.

Oakley, F., Kielhofner, G., Barris, R., & Reichler, R. K. (1986). The role checklist: Development and empirical assessment of reliability. Occupational Therapy Journal of Research, 6(3), 157-169.

Ottenbacher, K. J. (1987). Research: its importance to clinical practice in occupational therapy. American Journal of Occupational Therapy, 41(4), 213-215.

Ottenbacher, K. J., & Barrett, K. A. (1990). Statistical conclusion validity of rehabilitation research. A quantitative analysis. American Journal of Physical Medicine and Rehabilitation, 69(2), 102-107.

Ottenbacher, K. J., & Cusick, A. (1993). Discriminative versus evaluative assessment: Some observations on goal attainment scaling. American Journal of Occupational Therapy, 47(4), 349-354.

Ottenbacher, K. J., Tomchek, S. D. (1993). Reliability analysis in therapeutic research: Practice and procedures. American Journal of Occupational Therapy, 47(1), 10-16.

Pollock, N. (1993). Client-centred assessment. American Journal of Occupational Therapy, 47(4), 298-301.

Pollock, N., Baptiste, S., Law, M., McColl, M. A., Opzoomer, A., & Polatajko, H. (1990). Occupational performance measures: A review based on the guidelines for the client-centred practice of occupational therapy. Canadian Journal of Occupational Therapy, 57, 77-81.

Portney, L. G., & Watkins, M. P. (1993). Foundations of clinical research. Application to practice. Connecticut: Appleton & Lange.

Primavera, A., Novello, P., Finocchi, C., Canevari, E., & Corsello, L. (1990). Correlation between Mini-Mental State Examination and quantitative electroencephalography in senile dementia of Alzheimer type. Neuropsychobiology, 23, 74-78.

Puderbaugh, J. K., & Fisher, A. G. (1992). Assessment of motor and process skills in normal young children and children with dyspraxia. Occupational Therapy Journal of Research, 12(4), 195-216.

Reed, K. L. (1984). Models of practice in occupational therapy. Baltimore: Williams and Wilkins.

Reed, K., & Sanderson, S. R. (1980). Concepts in occupational therapy. Baltimore: Williams and Wilkins.

Reed, S. K. (1992). Cognition (3rd ed.). Pacific Grove: Brook/Cole Publishing Company.

Rogers, J. (1983). The study of human occupation. In Kielhofner, G. (Ed.), Health through occupation: Theory and practice in occupational therapy. Philadelphia: F.A. Davis.

Rothman, M. L., Hedrick, S., & Inui, T. (1989). The Sickness Impact Profile as a measure of the health status of noncognitively impaired nursing home residents. Medical Care, 27(3), S157-S167.

Sattler, J. J. (1992). Assessment of children (3rd.ed. Revised and Updated). San Diego: Jerome M. Sattler Publisher.

Schwamm, L. H., Van Dyke, C., Kiernan, R. J., Merrin, E. L., & Mueller, J. (1987). The Neurobehavioral Cognitive Status Examination: Comparison with the Cognitive Capacity Screening Examination and the Mini-Mental State Examination in a neurosurgical population. Annals of Internal Medicine, 107, 486-491.

Shah, S., & Cooper, B. (1993). Issues in the choice of activities of daily living assessment. Australian Occupational Therapy Journal, 40(2), 77-82.

Shepard, L. A. (1993). Evaluating test validity. In L. Darling-Hammond (Ed.), Review of Research in Education (Vol.19, pp.405-450). Washington, DC: American Educational Research Association.

Simon, H. A. (1979). Models of thought. New Haven: Yale University Press.

Smith, R. O. (1990). Administration and scoring manual. OT FACT (Occupational Therapy Functional Assessment Compilation Tool). Rocksille, Maryland: The American Occupational Therapy Association, Inc.

Smith, R. O. (1992). Guest editorial: The science of occupational therapy assessment. Occupational Therapy Journal of Research, 12(1), 3-15.

Smith, M. L., & Glass, G. (1987). Research and evaluation in education and the social sciences. N.J.: Prentice Hall.

Spencer, J. C. (1991). The physical environment and performance. In C. Christiansen and C. Baum (Eds.). Occupational therapy. Overcoming human performance deficits (Chapter 5, pp.126-140). Thorofare, NJ: SLACK Inc.

SPSS Inc. (1994). SPSS Professional Statistics 6.1. Chicago, IL: SPSS Inc.

Sternberg, R. J. (1986). Intelligence applied: Understanding and increasing your intellectual skills. San Diego: Harcourt Brace Jovanovich.

Sumsion, T. (1993). Client-centred practices: The true impact. Canadian Journal of Occupational Therapy, 60(1), 6-8.

Thorn, D. W., & Deitz, J. C. (1989). Examining content validity through the use of content experts. Occupational Therapy Journal of Research, 9(6), 334-345.

Tombaugh, T. N., & McIntyre, N. J. (1992). The Mini-Mental State Examination: A comprehensive review. Journal of American Geriatrics Society, 40, 922-935.

Townsend, E., Brintnell, S., & Staisey, N. (1990). Developing guidelines for client-centred occupational therapy practice. Canadian Journal of Occupational Therapy, 57(2), 69-76.

Trombly, C. (Ed.). (1988). Occupational therapy for physical dysfunction (3rd ed.). Baltimore: Williams & Wilkins.

Trombly, C. (1993). Anticipating the future: Assessment of occupational function. American Journal of Occupational Therapists, 47(3), 253-257.

Wellman, B. (1978). Applying social network analysis to the study of support. In B.H. Gottlieb (Ed.), Social networks and social support (pp. 171-200). Beverly Hill: Sage Publication.

Wood-Dauphinee, S. L., Opzoomer, M. A., Williams, J. I., Marchand, B., & Spitzer, W. O. (1988). Assessment of global function: The Reintegration to Normal Living Index. Archives of Physical Medicine and Rehabilitation, 69, 583-590.

World Health Organization. (1980). International classification of impairments, disabilities, and handicaps. Geneva, World Health Organization.

Yerxa, E. J. (1993). Occupational science: A new source of power for participants in occupational therapy. Occupational Science: Australia, 1(1), 3-10.

Yerxa, E. J., & Baum, S. (1986). Engagement in daily occupations and life satisfaction among people with spinal cord injuries. Occupational Therapy Journal of Research, 6(5), 271-283.

Yerxa, E. J., Burnett-Beaulieu, S., Stocking, S., & Azen, S. P. (1988). Development of the Satisfaction with Performance Scaled Questionnaire (SPSQ). American Journal of Occupational Therapy, 42(4), 215-221.

Yerxa, E. J., Clark, F., Frank, G., Jackson, J., Parham, D., Pierce, D., Stein, C., & Zemke, R. (1989). An introduction to occupational science, A foundation for occupational therapy in the 21st century. Occupational Therapy in Health Care, 6(4), 1-17.

APPENDIX I

Glossary of Terms

Model of Occupational Performance

ACTIVITY	A specific action, function or sphere of action that involves learning or doing by direct experience (Reed & Sanderson, 1980).
HOLISTIC	Perception of the client as a whole person; the overall state of health being interpreted as a result of a complex interaction of factors which includes mental, physical, sociocultural and spiritual components.
LEISURE	Components of life free from work and self-care activities.
OCCUPATIONAL PERFORMANCE	Activities carried out by the client in the areas of self-care, productivity and leisure influenced by environmental and societal factors (Reed & Sanderson, 1980).
OCCUPATIONAL THERAPY	The art and science which utilizes the analysis and application of activities specifically related to occupational performance in the areas of self-care, productivity and leisure. Through assessment, interpretation, and intervention, occupational therapists address problems impeding functional or adaptive behavior in persons whose occupational performance is impaired by illness or injury, emotional disorder, developmental disorder, social disadvantage, or the aging process. The purpose is to prevent disability; and to promote, maintain or restore occupational performance, health and spiritual well-being; furthermore, occupational therapy services can be directed through health, educational and social service systems.

PERFORMANCE COMPONENT - MENTAL	Mental - total emotional and intellectual response of an individual to the environment. Mentally, man has the capacity to reason and to adapt. Through interaction with the environment and with people, man develops a sense of self.... Man's responsibility for self evolves to provide the capacity to direct his own life.
PERFORMANCE COMPONENT - PHYSICAL	Physical - motor skills and sensory functions. Man satisfies needs or food, safety, sex and self-care. Through refined manual skills, he increases his potential to engage in activity.
PERFORMANCE COMPONENT - SOCIOCULTURAL	Sociocultural - dimension which describes the interpersonal relationships of a client with his family; and educational, ethnic and community background. Man's behavior patterns are determined by his set of beliefs, value system, developmental stage and life situation. His value system provides him with a blueprint for normative behavior.... During interaction with others, man is stimulated to experience emotions and feelings.
PERFORMANCE COMPONENT - SPIRITUAL	Spiritual - state of well-being; the force that permeates and give meaning to all life. Man is concerned with nature, the meaning of life and his purpose and place in the universe.
PRODUCTIVITY	Activities or tasks done to provide meaning and support to the self, family and society.
SELF-CARE	Activities or tasks done routinely to maintain the client's health and well-being in the environment.

Source:

Canadian Association of Occupational Therapists. (1991). Occupational therapy guidelines for client-centred practice. Toronto, ON: CAOT Publications.

APPENDIX II

Sample Items of the Role Checklist (Oakley et al., 1986)

The purpose of this checklist is to identify the major roles in your life. The checklist, which is divided into two parts, presents 10 roles and defines each one.

PART 1

Beside each role, indicate, by checking the appropriate column. If you performed the role in the past, if you presently perform the role, and if you plan to perform the role in the future. You may check more than one column for each role. For example, if you volunteered in the past, do not volunteer at present, but plan to in the future, you would check the past and future column.

Role:

STUDENT:
Attending school on a part-time or full-time basis.

WORKER:
Part-time or full-time paid employment.

VOLUNTEER:
Donating services, *at least once a week*, to a hospital, school, community, political campaign, and so forth.

CARE GIVER:
Responsibility, *at lease once a week*, for the care of someone such as a child, spouse, relative, or friend.

HOME MAINTAINER:
Responsibility, *at least once a week*, for the upkeeping of the home such as housecleaning or yardwork.

FRIEND:
Spending time or doing something, *at least once a week*, with a friend.

FAMILY MEMBER:
Spending time or doing something, *at least once a week*, with a family member such as a child, spouse, parent, or other relative.

RELIGIOUS PARTICIPANT:
Involvement, *at least once a week*, in groups or activities affiliated with one's religion (excluding worship).

HOBBYIST/AMATEUR:
Involvement, *at least once a week*, in a hobby or amateur activity such as sewing, playing a musical instrument, woodworking, sports, the theater, or participation in a club or team.

PARTICIPANT IN ORGANIZATIONS:
Involvement, *at east once a week*, in organizations such as the American Legion, National Organization for Women, Parents Without Partners, Weight Watchers, and so forth.

OTHER: ______________
A role not listed which you have performed, are presently performing, and/or plan to perform. Write the role on the line above and check the appropriate column(s).

APPENDIX III

Sample Items of the Satisfaction with Performance Scaled Questionnaire SPSQ (Yerxa et al., 1988)

During the last six months have you performed the following activity in such a way the you have felt happy, pleased, or contended with what you have done? In other words, how much of the time have you felt *satisfied* with the way you have done these activities?

I. Home Management

1. Scrape/stack dishes
2. Wash pots and pans
3. Remove/put away utensils/dishes in cupboards over sink/counters
4. Set/clear table
5. Load/unload washing machine
6. Dust high surfaces
7. Remove/put away utensils/dishes in cupboards under sink/counters
8. Use a floor mop
9. Make a bed
10. Use stove top elements
11. Put clothes on hangers
12. Clean a bathtub/shower
13. Reach high cupboards
14. Dispose of garbage
15. Sort clothes for washing
16. Open screw-top lids
17. Put clothes away in drawers/closet rod
18. Handle milk cartoon
19. Use a vacuum cleaner
20. Clean up counter/cooking surfaces
21. Get objects off top store shelves
22. Clean vegetables
23. Carry hot foods to table
24. Stir against resistance in a bowl

APPENDIX IV

Sample Items of the Klein-Bell ADL Scale
(Klein & Bell, 1979)

DRESSING

A. Obtaining clothing from bureau

1.	Grasp drawer	(1)
2.	Pull drawer open	(2)
3.	Reach into drawer	(2)
4.	Grasp clothes	(1)
5.	Shut drawer	(1)

B. Obtain clothing from closet

6.	Grasp clothing hung in closet	(1)
7.	Place clothes within reach for dressing	(1)

C. Socks

8.	Grasp sock	(1)
9.	Reach sock to L foot	(2)
10.	Reach sock to R foot	(2)
11.	Pull sock over R toes	(2)
12.	Pull sock over L toes	(2)
13.	Pull sock over R foot with heel to heel	(2)
14.	Pull sock over L foot with heel to heel	(2)
15.	Pull sock up to full extension on R leg	(2)
16.	Pull sock up to full extension on L leg	(2)

APPENDIX V

Sample Items of the Functional Independence Measure (Center for Functional Assessment Research, 1991)

Self Care

A. Eating
B. Grooming
C. Bathing
D. Dressing-Upper Body
E. Dressing-Lower Body
F. Toileting

Sphincter Control

G. Bladder Management
H. Bowel Management

Mobility

Transfer:
I. Bed, Chair, Wheelchair
J. Toilet
K. Tub, Shower

Locomotion

L. Walk/wheel Chair
M. Stairs

Communication

N. Comprehension
O. Expression

Social Cognition

P. Social Interaction
Q. Problem Solving
R. Memory

APPENDIX VI

Panel Review Questionnaire

NAME OF REVIEWER: ____________________________

DATE: _________________

INSTRUCTIONS TO THE REVIEWERS:

1. This questionnaire consists of three sections. The items in each section guide your evaluation of the COPM in terms of its test content, scoring structure, and test utility and procedure.

2. A five-point rating scale is used for each item ranging from Poor (1) to Excellent (5). Please circle the number corresponding to your evaluation for each item. Some of the items are accompanied with open-ended questions. Please provide brief comments, justifications, or suggestions wherever appropriate.

3. A brief review on the model of occupational performance is presented on page two of this questionnaire. Its purpose is to refresh your memory on the particulars of the model. If you are familiar with the model, please skip page two and begin the review of the COPM on page three of this questionnaire.

4. The questionnaire should not take more than one hour to complete. If there is any queries, please contact the investigators for clarification. Your cooperation is much appreciated.

Brief Review of the Model of Occupational Performance

The occupational performance model (CAOT, 1991) suggests a two-level domain of concerns which guide clinical practice of therapists. The higher level is occupational performance, which refers to the activities carried out by a person in the areas of self-care, productivity, and leisure. The lower level is performance components, which mean the basic building blocks of human functions such as abilities and skills, including mental, physical, sociocultural, and spiritual. According to the model,

> "...man has a need to be engaged. His engagement takes many forms and roles, each having a crucial effect on his quality of life. The culmination of man's experiences in this conceptual model is the formation of the individual's mental, physical, sociocultural and spiritual self. The essence of a healthy, functioning person is the balanced integration of these four performance components to provide a sense of well-being." (CAOT, 1991, p.6-17).

The person-environment-performance (PEP) framework developed by Christiansen (1991a) gives an excellent depiction of various factors that influence occupational performance. The higher order factors which are related to volition and control of behavior are: sense of competence, locus of control, and satisfaction. Other factors are grouped under the heading of "intrinsic enablers of performance" including psychological and cognitive, sensory and perceptual, neuromotor and physiological factors.

Reference:

Canadian Association of Occupational Therapists. (1991). Occupational therapy guidelines for client-centred practice. Toronto, ON: CAOT Publications.

Christiansen, C. (1991a). Occupational therapy. Intervention of life performance. In C. Christiansen and C. Baum (Eds.). Occupational therapy. Overcoming human performance deficits (Chapter 1, pp.1-43). Thorofare, NJ: SLACK Inc.

CONTENT REVIEW

1.1 How well does the COPM reflect the self-care performance of a person being assessed?

Rating:

1	2	3	4	5
Poor	Fair	Good	Very Good	Excellent

Which part of the COPM do you think assesses the self-care performance?

1.2 How well does the COPM reflect the productivity performance of a person being assessed?

Rating:

1	2	3	4	5
Poor	Fair	Good	Very Good	Excellent

Which part of the COPM do you think assesses the productivity performance?

1.3 How well does the COPM reflect the leisure performance of a person being assessed?

Rating:

1	2	3	4	5
Poor	Fair	Good	Very Good	Excellent

Which part of the COPM do you think assesses the leisure performance?

1.4 In general, how well does the COPM reflect the occupational performance of a person being assessed?

Rating:

1	2	3	4	5
Poor	Fair	Good	Very Good	Excellent

1.5 How well does the COPM reflect the mental performance component, such as emotional and intellectual functions, of a person being assessed?

Rating:

1	2	3	4	5
Poor	Fair	Good	Very Good	Excellent

Which step(s) or process(es) reflect the evaluation of mental abilities of a person?

1.6 How well does the COPM reflect the physical performance component, such as motor skills and sensory functions, of a person being assessed?

Rating:

1	2	3	4	5
Poor	Fair	Good	Very Good	Excellent

Which step(s) or process(es) reflect the evaluation of physical capabilities of a person?

____

____

1.7 How well does the COPM reflect the sociocultural performance component, such as set of beliefs, value system, developmental stage, and life situation, of a person being assessed?

Rating:

1	2	3	4	5
Poor	Fair	Good	Very Good	Excellent

Which step(s) or process(es) reflect the evaluation of sociocultural performance component of a person?

____

____

1.8 How well does the COPM reflect spiritual performance component, such as meaning of life and his purpose and place in the universe, of a person being assessed?

Rating:

1	2	3	4	5
Poor	Fair	Good	Very Good	Excellent

Which step(s) or process(es) reflect the evaluation of spiritual performance component of a person?

1.9 In general, how well does the COPM reflect the performance components of a person being assessed?

Rating:

1	2	3	4	5
Poor	Fair	Good	Very Good	Excellent

1.10 On the whole, how well does the COPM completely represent the philosophy, rationale, and frame of reference of the model of occupational performance? You may want to refer to page one of this Questionnaire for a brief review of the model (CAOT, 1991).

Rating:

1	2	3	4	5
Poor	Fair	Good	Very Good	Excellent

Please briefly justify your choice.

1.11 On the whole, to what degree is the content of COPM, such as problem definition and importance, perceived performance, and satisfaction ratings, relevant to the occupational performance of a person being assessed?

Rating:

1	2	3	4	5
Poor	Fair	Good	Very Good	Excellent

Please briefly justify your choice.

STRUCTURAL REVIEW

2.1 How well does the ten-point rating scale (used in COPM) quantify the degree of importance of an activity in a person's life?

Rating:

1	2	3	4	5
Poor	Fair	Good	Very Good	Excellent

Suggestions for changes, if any:

____

____

2.2 How well does the ten-point rating scale (used in COPM) quantify a person's subjective evaluation of his/her current performance?

Rating:

1	2	3	4	5
Poor	Fair	Good	Very Good	Excellent

Suggestions for changes, if any:

____

____

2.3 How does the ten-point rating scale (used in COPM) quantify a person's subjective evaluation of his/her satisfaction with current performance?

Rating:

1	2	3	4	5
Poor	Fair	Good	Very Good	Excellent

Suggestions for changes, if any:

____

____

2.4 To what degree does the computation of the Performance scores (multiply Importance by Performance) for each identified activity reflect the structural relations between a person's occupational performance and performance components?

Rating:

1	2	3	4	5
Poor	Fair	Good	Very Good	Excellent

Suggestions for changes, if any:

____

____

2.5 To what degree does the computation of the Satisfaction scores (multiple Importance by Satisfaction) for each identified activity reflect the structural relations between a person's occupational performance and performance components?

Rating:

1	2	3	4	5
Poor	Fair	Good	Very Good	Excellent

Suggestions for changes, if any:

2.6 To what degree does the computation of the Performance subscale score (average of five item Performance scores) reflect the occupational performance of a person as a product of the interactions of his/her performance components?

Rating:

1	2	3	4	5
Poor	Fair	Good	Very Good	Excellent

Suggestions for changes, if any:

2.7 To what degree does the computation of the Satisfaction subscale score (average of five item Satisfaction scores) reflect the occupational performance of a person as a product of the interactions of his/her performance components?

Rating:

1	2	3	4	5
Poor	Fair	Good	Very Good	Excellent

Suggestions for changes, if any:

2.8 How well does the Change in Performance and Change in Satisfaction scores reflect the outcomes of occupational therapy intervention in terms of occupational performance of a person?

Rating:

1	2	3	4	5
Poor	Fair	Good	Very Good	Excellent

UTILITY AND PROCEDURAL REVIEW

3.1 Are guidelines in the COPM Test Manual set forth for potential examiners in terms of training, credentials, and/or theoretical background?

Rating:

1	2	3	4	5
Poor	Fair	Good	Very Good	Excellent

3.2 Is the training process provided in the COPM Test Manual explicitly presented or offered to insure inter-rater reliability?

Rating:

1	2	3	4	5
Poor	Fair	Good	Very Good	Excellent

3.3 Is the complexity of administration of the COPM congruent with the levels of required training and potential clinical decisions?

Rating:

1	2	3	4	5
Poor	Fair	Good	Very Good	Excellent

3.4 Are the administrative procedures described in the COPM Test Manual clear and precise?

Rating:

1	2	3	4	5
Poor	Fair	Good	Very Good	Excellent

3.5 Does the administration of the COPM take a reasonable amount of time to complete?

Rating:

1	2	3	4	5
Poor	Fair	Good	Very Good	Excellent

3.6 Is all the necessary equipment for administering the COPM included in the Test Manual? Is the equipment safe, durable, manageable, and replaceable?

Rating:

1	2	3	4	5
Poor	Fair	Good	Very Good	Excellent

3.7 Is the COPM designed and standardized for the population on which you wish to use it?

Rating:

1	2	3	4	5
Poor	Fair	Good	Very Good	Excellent

3.8 Is adequate information available in the Test Manual and/or related literature regarding the method of standardization, the standardization sample, and/or the field testing of the COPM?

Rating:

1	2	3	4	5
Poor	Fair	Good	Very Good	Excellent

3.9 Is evidence available in the Test Manual and/or related literature on the validity of the interpretation and use of the COPM?

Rating:

1	2	3	4	5
Poor	Fair	Good	Very Good	Excellent

3.10 Is evidence available in the Test Manual and/or related literature on the internal consistency, test-retest, and/or inter-rater reliability of the COPM?

Rating:

1	2	3	4	5
Poor	Fair	Good	Very Good	Excellent

3.11 Do the sub-scale and total scores of the COPM reasonably lend themselves to interpretation?

Rating:

1	2	3	4	5
Poor	Fair	Good	Very Good	Excellent

3.12 Is the COPM applicable to the assessment of client's occupational performance across all developmental levels?

Rating:

1	2	3	4	5
Poor	Fair	Good	Very Good	Excellent

3.13 Is the COPM applicable to the assessment of client's occupational performance with all disability groups?

Rating:

1	2	3	4	5
Poor	Fair	Good	Very Good	Excellent

3.14 Is the Test Manual of the COPM complete, well organized, and easy to use?

Rating:

1	2	3	4	5
Poor	Fair	Good	Very Good	Excellent

3.15 Is COPM's test format organized, logical, and appealing?

Rating:

1	2	3	4	5
Poor	Fair	Good	Very Good	Excellent

3.16 Would you consider the COPM as a cost effective assessment?

Rating:

1	2	3	4	5
Poor	Fair	Good	Very Good	Excellent

3.17 Is the COPM of a quality that would be used and respected by other health professionals?

Rating:

1	2	3	4	5
Poor	Fair	Good	Very Good	Excellent

3.18 Were references and corollary studies cited or conducted and reported in the Test Manual of the COPM?

Rating:

1	2	3	4	5
Poor	Fair	Good	Very Good	Excellent

3.19 Does the COPM represent the highest state of art in the measurement of occupational performance of individuals?

Rating:

1	2	3	4	5
Poor	Fair	Good	Very Good	Excellent

THANK YOU FOR YOUR PARTICIPATION !!!

APPENDIX VII

Educational Session

LOCATION: Glenrose Rehabilitation Hospital, Edmonton
DATE: September 8th, 1993
TIME: 8:00am to 10:00am

Purposes of the Session:

1. Describe the operational definitions pertinent to the validation study.
2. Standardize the procedures of administering the COPM.
3. Receive the data base sheet.
4. Confirm allocation of clients to participating therapists.

Operational Definitions:

Semi-structured Interview

It is defined in terms of the type of questions that are asked by the participating therapists when he/she administers the COPM. Semi-structured questions have no choices from which the respondent selects an answer. Questions are phrased to allow for individual responses. It is an open-ended question but fairly specific in its intent (McMillan & Schumacher, 1989).

Occupational Performance

Activities carried out by the client in the areas of self-care, productivity, and leisure influenced by environmental and societal factors (CAOT, 1991).

Self-Care

Activities or tasks that are done routinely to maintain the client's health and well-being in the environment (CAOT, 1991).

Productivity

Activities or tasks that are done to provide meaning and support to the self, family, and society (CAOT, 1991).

Leisure

They are the activities or components of life that are free from work and self-care activities (CAOT, 1991).

Disability

It is any restriction or lack (resulting from an impairment) of ability to perform an activity in the manner or within the range considered normal for a human being (WHO, 1980). Two digit categories of disability are used in the subject's data base sheet.

COPM Administration Procedures

Assessment Schedule:

Initial Assessment:
1) orthopedic client subjects: within three days;
2) stroke client subjects: within the first two weeks;
after the reception of their requisition for occupational therapy consultation.

Pre-discharge Assessment:
1) orthopedic client subjects: within three days;
2) stroke client subjects: within one week;
before they are discharged (or with respect to a tentative date of discharge) to home, or other placements, from the hospital.

Consent from Subjects:

- participating therapist:
    1) explain the purposes of the project to the client subject;
    2) give the Client Information Sheet to the client subject;
    3) obtain consent from the client subject by signing the Consent.

Date Base:

- Participating therapist fills out the set data base form (see attached form).

ITEMS: Demographic data: Name of client, Sex, Age, ID No.
Ethnicity
Environment (home situation)
Support System
Length of Stay in Hospital (no. of days)
O.T. Admission and Discharge Dates

COPM Cover Page:

- No additional instructions. Follow procedures as in the Test Manual.

COPM Step 1: Problem Definition:

- Follow instructions as in the Test Manual.

- Therapist is encouraged not to use the list of activities (Appendix C) in the Test Manual.

- Therapist should encourage the client subject to identify as many activities that he/she has problems with as possible.

COPM Step 2: Weighting the Problems:

- Follow instructions as in the Test Manual.

COPM Step 3: Scoring:

- Follow instructions as in the Test Manual.

COMP Step 4: Re-assessment:

- Same as in Step 3.

COPM Step 5: Follow-up:

- This step is redundant because re-assessment is done at the pre-discharge stage.

APPENDIX VIII

CONSENT FORM

Project Title: Validation of the Canadian Occupational Performance Measure

Investigators:
Chetwyn Chan, University of Alberta
Annette Magnan & Lorian Kennedy, Occupational Therapy
Tom Maguire, PhD., University of Alberta
______________________, Occupational Therapy

This research study will evaluate a new test in occupational therapy. This test measures clients' performance and improvement.

I agree to participate in two interviews using the Canadian Occupational Performance Measure (COPM). On each occasion, I also agree to answer two questionnaires. One questionnaire asks about my life roles and the other measures my satisfaction with the way I can do things. All the tests and questionnaires will take a total of two hours. I also agree to participate in the screening exercise and to let the investigators obtain needed information from my medical records.

The study carries no risks to me. There will be no direct benefits for me. My name will not appear in any documents or reports. I can refuse to answer any items in the interviews and the questionnaires. All information collected in this study will be kept confidential.

I am free to withdraw my consent and stop participating at any time. This will not affect my present or future care. I HAVE BEEN GIVEN THE CHANCE TO ASK QUESTIONS. I AM SATISFIED THAT ALL MY QUESTIONS HAVE BEEN ANSWERED. My signature means:

1) I have read this form
2) I understand my involvement in the study
3) I voluntarily agree to participate.

I will be given a copy of this consent form. If I have any questions concerning the study I can contact Chetwyn Chan, 492-3762; or Annette Magnan, 471-2262 Ext.2362.

Name of Participant

Signature of Participant — Date

Signature of Investigator — Date

Signature of Witness — Date

APPENDIX IX

CLIENT INFORMATION SHEET

Validation of the Canadian Occupational Performance Measure

The Canadian Occupational Performance Measure (COPM) is a new test developed by the Canadian Association of Occupational Therapists and Department of Health and Welfare. The COPM is intended to measure skills in the areas of self-care, productivity, and leisure. This research project will look at how good the test is for measuring these things.

In this study, an O.T. will give you the COPM twice. The first COPM will be at the beginning of your stay. The second will be just before your discharge from the hospital. Besides the COPM, you will be asked to fill out two other questionnaires: the Role Checklist and the Satisfaction with Performance Scaled questionnaire. These are both paper and pencil measures. If you need help with these, your therapist will be glad to assist you. For the first session, it will take about one hour to finish all the tests after you complete a screening exercise. Less time is needed for the second session.

You may be asked to participate in a retest. The purpose of the retest is to see how much the COPM results change within a few days. In the retest, you will be asked do the COPM again. Also, the therapist will ask you a few questions about how you chose your answers.

The data collected from the tests is for research purposes. If you would like, the researcher can release the results to your treating therapist. This may be helpful for your therapy.

Thank you for your participation.

Investigators,

Chetwyn Chan
Annette Magnan
Lorian Kennedy
Tom Maguire

APPENDIX X

Validation of the Canadian Occupational Performance Measure
Client Data Base Sheet

Disability Classification (use 2-digits WHO System) ____________
____
____

Ethnicity / Culture:

White ______
Black ______
Native ______

Asian ______
Hispanic ______
Other ______
Specific ____________________

Home Situation:

Comment: ____
____
____

Social Situation:

Marital Status ______
Lives Alone ______ Lives with Others ______
Specify ________________
Support Available at Discharge Yes______ No______
Specify____________________

Management Statistics:

Date of Initial Ass. ________________
Date of Pre-discharge Ass. ________________
Date of Discharge ________________
Date of Admission ________________
Length of Stay in Hosp. ___________ Days

APPENDIX XI

Activity Reference List
(Adapted from Law et al., 1991, p.31-34)

1. Self Care

11 Personal Care

01 getting clothes from closets and drawers (1)
02 managing fasteners (3)
03 removing clothing (1)
04 opening containers
05 pouring liquids
06 toiletting (5)
07 washing self
08 brushing teeth
09 maintaining nails (3)
10 shaving (1)
11 applying makeup (1)
12 engaging in sexual activity
13 taking medication
14 eating (4)
15 dressing upper and lower limb garments (19)
16 putting on and off shoes and socks (10)

12 Functional Mobility

01 turning in bed (6)
02 transferring from wheelchair (2)
03 getting in and out of tub (14)
04 climbing up and down stairs (12)
05 getting in and out of cars, taxi (18)
06 walking (21)
07 getting in and out of bed (12)
08 getting in and out of toilet (1)

13 Community Management

01 driving car (11)
02 taking public transit (3)
03 using telephone (1)
04 arranging for services, e.g. plumber
05 making appointments
06 budgetting
07 paying bills (2)
08 handling money, making cheques
09 filing taxes
10 shopping (17)

2. Productivity

21 Paid/Unpaid Work

01 job searching (1)
02 preparing resume
03 arranging for interview
04 attending interviews (1)
05 selecting a job
06 working expected hours (11)
07 learning new tasks (1)
08 managing responsibilities (1)
09 relating to co-workers
10 adhering to safety regulations
11 being punctual
12 dealing with problems and conflicts (1)

22 Household Management

01 buying groceries (2)
02 planning meals
03 following a recipe
04 preparing foods (13)
05 clearing the table
06 doing the dishes (2)
07 cleaning the kitchen (2)
08 dusting and polishing furniture (6)
09 sweeping the floor (2)
10 vacuuming (5)
11 disposing of garbage (1)
12 cleaning bathrooms
13 doing laundry (14)
14 ironing (1)
15 minor repairs, e.g. replacing light bulbs
16 caring for children
17 making beds (6)

23 Play / School

01 exploring and manipulating objects
02 climbing, throwing, running, jumping
03 playing with peers
04 playing cooperative games
05 sharing
06 taking turns
07 engaging in sports
08 role-playing, pretending
09 creative expression
10 separating from parents
11 drawing, cutting, pasting
12 printing / writing (2)
13 subtracting, multiplying etc
14 verbalizing reporting
15 remembering
16 completing homework
17 note-taking

3. Leisure

31 Quiet Recreation

01	listening to music
02	watching television (1)
03	reading books, newspapers, magazines (8)
04	knitting, sewing, crocheting (3)
05	working on hobbies, collections (7)
06	creative arts, media
07	playing cards, board games (1)
08	gardening (2)

32 Active Recreation

01	participating in sports (10)
02	caring for a pet (3)
03	going to movies, theatre (4)
04	dining out (1)
05	going to bars, nightclubs
06	driving, sightseeing
07	going to parks, beaches, playgrounds (4)
08	attending religious services (2)
09	attending courses (3)
10	visiting museums or libraries
11	travelling (3)
12	attending sports events (2)
13	dancing (1)

33 Socialization

- 01 visiting with friends / family (9)
- 02 talking on the telephone (1)
- 03 attending parties (2)
- 04 hosting parties
- 05 writing letters (2)
- 06 planning social engagement
- 07 attending group functions (2)

Note. Figures in parentheses are frequencies of the activity identified by the clients as difficult to perform in this study. The total number of activities identified is 327.